L LIBRARY
OF MICH.

AUG 24 1936

VOL. XXVII MAY, 1936 NO. 5

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

CONTENTS

Attitude Difference between College Classes: A Summary and Criticism 321
STEPHEN M. COREY

Studies of Jewish and Non-Jewish Intelligence 331
MOSHE BRILL

Introduction to Inverted Factor Analysis, with Some Applications to Studies in Orexis 353
W. STEPHENSON

The Scoring of Matching-tests with Unequal Series of Items 368
DWIGHT W. CHAPMAN

Nomographs Concerning the Spearman-Brown Formula and Related Functions 371
J. NORMAN ARNOLD AND JACK W. DUNLAP

A Factor Analysis of Personality Traits to Underlie Character Education 375
C. H. McCLOY

Note on the Comparative Efficacy of Lecture and Socialized Recitation Method vs. Group Study Method 388
WILLIAM H. BARNARD

Book Reviews 391

$6.00 per Year · Published Monthly September to May

WARWICK & YORK, INC.

BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as Second Class Matter at York, Pa.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

Established 1910

EDITORS

JACK W. DUNLAP
Fordham University

HAROLD E. JONES
University of California

PERCIVAL M. SYMONDS
Teachers College,
Columbia University

H. E. BUCHHOLZ, *Managing Editor*

THE price of the Journal is $6.00 a year in the United States; $6.40 in foreign countries. Part-year subscriptions are 90 cents for each number ordered. Back volumes are $7.00 each; back issues are $1.10 each except when more than five years old, and then $1.20 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Unsolicited manuscripts should be accompanied with return postage. Manuscripts, books and other materials for review, and correspondence regarding editorial and business matters should be addressed to the Publishers.

WARWICK AND YORK · *Publishers* · BALTIMORE, MD.

THE ACTIVITY MOVEMENT

By CLYDE HISSONG

In an attempt to overcome the weaknesses of the traditional school organization many progressive schools have developed new programs. These programs are so similar in character that collectively the changes have been referred to as the activity movement. This movement has claimed the center of the educational stage for a length of time sufficient to have engendered widespread interest in its outcomes and in its basic philosophy.

In Doctor Hissong's study an attempt has been made to discover the principles underlying the present activity movement, to determine the influence of traditional concepts in shaping the trends of the movement, and to see if in the light of the present knowledge of the child and his relation to his environment the movement rests upon a justifiable basis.

$2.00 plus 10¢ postage

WARWICK AND YORK

Publishers BALTIMORE

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXVII May, 1936 Number 5

ATTITUDE DIFFERENCES BETWEEN COLLEGE CLASSES: A SUMMARY AND CRITICISM

STEPHEN M. COREY

University of Nebraska

A number of psychologists and sociologists have interested themselves in the attempt to measure the variations which occur from year to year in the attitudes of college students. This is an important problem. Most educators assume, when they are not busy stating so explicitly, that at all levels of formal training there should be shifts in attitudes, points of view, and opinions. An inspection of some sixty-eight private and public college bulletins revealed a great number of changes of this sort which it was hoped might occur in the students who were exposed to the influence of this or that institution of higher learning.* Thirty-four of these bulletins, for example, mentioned religious attitudes, thirty-two stressed the importance of certain attitudes toward intellectual pursuits, thirty described attitudes toward social institutions, and nineteen, certain desirable attitudes toward citizenship.

The technique which has been customarily employed to determine whether these attitudinal changes have taken place involves, first, the administration of a questionnaire to a large group of college students including freshmen, sophomores, juniors, and seniors. The results of this inquiry are then tabulated and inferences made regarding changes in attitudes in terms of the different reactions of students in the four college classes. If an investigator wished to know, for example, whether a college education had any effect upon student attitudes toward walking on the grass, he would, more than likely, go through these four steps:

1. Construct an attitude scale, in all probability according to the Thurstone technique.

* From an unpublished study by President Erland N. Nelson of Dana College.

2. Administer the scale to part of the college population.

3. Score separately the scales of freshman, sophomore, junior, and senior groups.

4. Hunt for "statistically significant differences" to justify inferences regarding the effect of a college education upon student attitudes toward walking on the grass.

THE SUMMARY

The writer has had his attention drawn to a considerable number of studies of this general type. Some of them are reviewed below in the chronological order of their publication.

Symonds[22] administered one hundred true-false questions alleged by five persons to measure different degrees of liberalism to students in Hawaii beginning with grade eight and including some individuals in all four undergraduate classes at the University of Hawaii. After scoring the papers from the different grade levels separately and finding few significant differences the author concluded, expressing surprise, " . . . that strictly school education has been able to make no change in the above attitudes beyond grade VIII." (p. 319) The inference is rather plain that had differences appeared they would have been explained, in part at least, as the result of formal education.

Jones[15] gave a questionnaire of twenty-five statements measuring attitudes toward (*a*) national and social optimism, (*b*) labor problems and economic status generally, (*c*) discipline, (*d*) social life and conventions, and (*e*) religion, to four hundred eighteen college students and tabulated scores for the various classes. After devoting con siderable space to a discussion of the differences between senior an freshman scores, his conclusion to this aspect of the study was:

> The most significant general educational implication is the slight effect that a college training has had on the real opinions of our students who are now seniors in the Arts and Science College. (p. 431)

In a somewhat similar study reported by Bain[1] a questionnaire was given to students at the University of Washington dealing primarily with liberalism *versus* conservatism in regard to religious beliefs. After the statement, "It is commonly believed that a college education tends to liberalize the religious views of students," Bain concluded, basing his inference on a comparison of scores made by members of the different college courses, that " . . . on all except a few questions the seniors were more unorthodox than the other boys by substantial percentages." (p. 764)

In their application of the scale measuring attitude toward the church, Thurstone and Chave[23] analyzed scores made by students in the four undergraduate classes at the University of Chicago. No statistically significant differences were apparent; so the authors concluded:

> The four undergraduate classes do not show any distinct trend to become more in favor of or more against the church as they progress through college. (p. 69)

Actually, of course, no class was measured as it progressed through college. The scale was given to four different classes, and the similarity of scores caused the authors to imply that attendance at college had little or no effect upon the group of religious attitudes measured.

Using much the same general technique, Willoughby[25] summarized questionnaire data resulting from an inquiry put to about eight hundred Leland Stanford students and dealing with attitudes toward a number of matters of current interest as follows:

> There appear to be traces of a systematic shift of opinion with longer academic experience, particularly in the direction of liberalism. (p. 169)

The inference that the "longer academic experience" was responsible for at least some of the increase in liberalism is apparent.

Garrison and Mann[11] administered a twenty-five statement questionnaire dealing with economic, legal, religious, civil and social issues to some two hundred fifty students from different college classes at North Carolina State College. The opinions of freshmen, sophomores, juniors, and seniors were analyzed and inferences such as these drawn:

> As to average certainty of opinion, it appears to be little affected by college courses. (p. 176)
>
> The tendency on the part of the upperclassmen to be more cautious and less suggestible, along with a greater amount of liberalism present in their attitudes on religious and social issues, would seem to point out some effect of training. (p. 176)

In this study the authors were quite explicit in their belief that attitudinal differences between college classes resulted in part from exposure to the influences of formal college education.

Droba,[6] who has published a number of papers on attitude measurement, administered his own "Attitude Toward War" scale to one thousand college students and made this deduction after an analysis of the scores of individuals in different college classes:

A general trend toward pacifism may be noted in the successive mean scores of the first three classes and the graduate class. This is an indication that education is influential in shaping the attitudes of students in the direction of pacifism. (p. 142)

The same author[7] reporting the standardization data for his scale, published results for the four college classes as well as the graduate college. Again basing his deductions on observed differences between scores made by freshmen, sophomores, juniors, and seniors, he concluded:

These results tend to indicate that education has something to do with the formation and modification of attitudes toward war and peace. (p. 108)

Education appears to develop pacifism as measured by the scale. (p. 110)

Again, in a study of the effect of education upon Negro attitudes, Droba[5] administered the Hinckley scale to small groups of juniors, seniors, and graduate students at Ohio State University. In view of the differences in attitude scale scores between these groups, he concluded:

Education at the Ohio State University, as indicated by a change in attitudes from year to year, tends to develop favorable attitudes toward the Negro. (p. 141)

In each of these three studies Droba employed the same reasoning. He found differences between the scores made by students at various levels of collegiate or university instruction and inferred that these differences were the result of continued exposure to higher education.

Katz and Allport[16] in their comprehensive study of students' attitudes attempted to measure beliefs concerning the Deity in relation to class in college. Seven statements about God were reacted to by approximately fourteen hundred Syracuse students, and comparisons made of the different opinions expressed by freshmen, sophomores, juniors and seniors. After calling attention to the two assumptions involved in any such attempt to infer development from different class reactions; namely, that each successive class entering college represents the same selection from the population at large, and that the religious viewpoints of society at large do not shift materially during a four-year interval, these authors concluded:

We can say therefore that the juniors and seniors differed from the freshmen and sophomores, not in their modal position, but in having a greater proportion of their number who stressed a liberal impersonal notion of the Deity rather than the old personal and orthodox view. And in so far as this difference indicates a change in view due to residence in college, it may be said

that the influence of a college training works slightly in the direction of unorthodoxy. (p. 269)

Garrison and Burch[10] asked one hundred sixty-three students to endorse certain of thirty-five statements taken from the Social Science Research Council's racial attitude questionnaire. These students were from all four classes at North Carolina State College, and an analysis of their responses led the authors to conclude, after calling attention to the similarity between several groups:

There is indication, however, that age, home environment, and degrees of college training are to some extent factors in determining differentiated racial attitudes. The seniors, for instance, show a somewhat more lenient attitude than do members of the lower classes. (p. 234)

Somewhat in opposition to Droba's conclusions[6,7] yet using the same statistical technique, Sowards[20] decided that education has slight, if any, effect upon war attitudes. She administered the Droba test to high-school seniors as well as college freshmen and seniors and concluded:

If college education modifies a person's attitude toward war, the modification is not very marked. (p. 329)

Again, after finding no significant difference between the scores of high-school seniors and freshmen in college, the conclusion is that " . . . education seems to exert little influence in determining or modifying war attitudes." (p. 330) It should be noted, however, that Miss Soward employed the same type of reasoning as did Droba. She interpreted the absence of differences between attitude-scale scores achieved by pupils at different educational levels as evidence of the ineffectiveness of formal education to bring about changes of this sort.

Dudycha[8] asked that a list of twenty-five propositions concerning evolution be marked *A*, *B*, *C*, *D*, or *E* according to the degree of their acceptability to seven hundred fifty freshmen and three hundred four seniors in six midwestern colleges. After finding a number of differences between the two groups he concluded:

Since the tests were administered near the close of the students' senior year their judgments were influenced by four years of college instruction. (p. 85)

The implication that at least some of the difference in attitudes between the two classes was due to this college training, inasmuch as attention was called to the fact that the freshmen had just started in school, is apparent.

Johnson[14] constructed an "Attitude Toward the Negro" scale after the Thurstone-Chave technique and administered it to some three thousand five hundred subjects, about eight hundred of whom were negroes. His analysis of the reactions characteristic of different college classes led him to conclude that "racial attitudes vary slightly according to education." (p. 25) In other words, slight differences in the attitudes of persons of varying educational status toward the Negro can be explained in part in terms of continued schooling.

Boldt and Stroud[2] administered the Harper Test to about seven hundred students in one of the Kansas State Teachers Colleges. The scores indicated an increase in liberality from the freshman year up through the graduate school. The increase was regular and consistent, and the authors concluded:

> The results of this investigation indicate that the attitudes of the college students tested became more liberal, as a result of their training, toward the issues involved in the test. (p. 619)

Miss Kellar[17] studied the attitude of students majoring in home economics toward "caring for children" and "meal preparation." She gave attitude scales to freshmen, sophomores, juniors, and seniors " . . . to see if their attitude toward the activities became more favorable as the students advance in college." Although it is evident that the attitude scales were not administered to students as they "advance in college" but rather to four different groups in four different classes, Miss Kellar concluded:

> There is no progression for attitude toward caring for children, but it is seen that the seniors like to prepare meals better than the underclassmen do. (p. 58)

Symington[21] in his investigation of religious liberals and conservatives likewise studied separately the scores of different educational groups on a "Test of Religious Thinking." After making frequent references to the "close relation of academic training and liberal religious thinking," he concluded:

> Those who are more advanced in school or college are more liberal in their religious thinking. (p. 43)

Little doubt is left in the reader's mind as to the author's belief in the causal relationship between these factors.

THE CRITICISM

The methodological limitation to which all* these studies are subject is that the conclusions do not take into account the following facts:

* To the best of the writer's knowledge there are only two studies, of the type

1. Students drop out of college in large numbers from the freshman to the senior year.

2. This mortality is selective on a number of important bases.

With regard to the first oversight, enrolment statistics for American colleges and universities indicate that in general about one freshman in three is ever graduated.* The mortality, while particularly marked during the freshman year, continues throughout the four years of college. The selective bases for this elimination have been made apparent in a great number of investigations.†

The bearing which facts such as the foregoing have upon conclusions to the studies summarized in this paper would need no elaboration were they not so persistently and consistently overlooked. To conclude, from a comparison of scores made by freshmen, sophomores, juniors, and seniors, that education either has or has not been effective in engendering or qualifying certain attitudes, is to overlook the very obvious fact that the differences between these scores, or for that matter, the absence of differences, might quite possibly be due to selection brought about by academic mortality. It would be just as fallacious to infer from differences between the psychological test scores of college freshmen, sophomores, juniors, and seniors that prolonged education resulted in marked and significant increases in intelligence. Education may have nothing to do with the matter, or, equally enlightening, actual attitudinal differences brought about by education might not be apparent because of academic mortality.

For example, let us assume that the following hypothetical scores have been achieved by students on attitude scale "*A*":

	Mean score	Standard deviation
Freshmen	6	.11
Sophomores	8	.11
Juniors	10	.12
Seniors	12	.13

being considered, which mention the possibility of selection being a factor in bringing about variations in attitudes as evidenced by individuals of differing educational status. These are, first, that reported by Harper in 1927[12] who claimed that his correlation of .52 between liberality and education beyond the eighth grade might be due to training or selection; and second, Kolstad's monograph *Opinions on International Problems*,[18] in which much the same admission was made.

* For a typical study see Edgerton and Toops.[9]

† For a bibliography of many studies of this type see Cowley.[4]

All these differences are statistically significant, and the customary interpretation would be that continued education or training brought about attitudinal changes. However, as has been noted above, this conclusion overlooks the facts of academic mortality. If the attitude being studied is correlated with student persistence, and this can only be determined by comparison of scores made by students remaining in school with those made by students dropping out, the differences might all be due to such selection. Conceivably the very direction of the apparent change might be reversed because of the attitude scores of students not returning to school.

To cite another hypothetical case—and the writer again apologizes for an elaboration of the obvious on the grounds that the matter apparently is not obvious—consider the following scores on attitude scale "*B*":

	Mean score	Standard deviation
Freshmen	8	.10
Sophomores	8	.10
Juniors	8	.10
Seniors	8	.10

Here the superficial conclusion might be that college training had no effect upon attitude "*B*," when, as a matter of fact, if this attitude is one of the factors related to student mortality, the mean scores given above tell us nothing about the manner in which it is affected by education. The attitude in question might have been strengthened, weakened, or not affected at all.

That group scores on attitude scales may be affected by academic mortality receives support from the following observations. A number of writers have reported that intelligence and "liberalism" are positively correlated.[3,13,15,18,19,21,24] In one instance[21] the correlation reported was as high as +.55. Most attitude scales do permit, in general, of reactions which involve the trait conservatism-liberalism. Illustrations from the Thurstone series are scales measuring attitudes toward militarism-pacifism, prohibition, birth control, treatment of criminals, God, Sunday observance, and a number of others. This being true, and keeping in mind the rather significant bearing of intelligence upon student mortality,[4,9] any attitude involving liberalism-conservatism, because of the relationship of this trait to intelligence, probably is more or less affected by persistence in college.

The writer has tabulated the results of one of Thurstone's attitude scales in such a way as to bring out the effect of student mortality.

Table I is based upon scores made by one hundred fifty-one freshman students who took the Thurstone-Chave "Attitude Toward the Church" scale in September, 1933. The mean score for this group is given in line one. Line two gives the mean freshman attitude test scores for those students who returned as sophomores in September, 1934. Similarly, line three gives the mean freshman attitude score for the sixty-six students who returned as juniors in September, 1935.

TABLE I.—COLLEGE MORTALITY AND CHANGES IN ATTITUDES

	No.	MN score	SD mean
1. Freshman score of all students September, 1933.....	151	2.54	.09
2. Freshman score of students returning September, 1934...	115	2.59	.11
3. Freshman score of students returning September, 1935...	66	2.64	.17

None of the differences between these scores is significant statistically, yet differences do occur and they indicate a "trend"—that is, for each successive group the attitude becomes more "liberal." At the present time it is impossible to determine the freshman scores of the students returning for their senior year. Were the "trend" to continue, the difference between the senior and freshman groups might be significant, altogether because of selective mortality. The scales were all given at the beginning of the students' first year at the University. At that time, of course, their opinions were unaffected by college training.

The popularity of the method of measuring changes in student attitudes by administering attitude scales to individuals with different degrees of academic training lies, very likely, in its convenience. It is astonishingly easy to administer the attitude scales, and the general availability of college students of different classes makes this type of investigation very enticing. The only technique which will give *valid* data involves the repeated testing of the *same* students as well as repeated testing of other young people not in college. This is a time-consuming process calling for a four- or five-year program of testing. In the opinion of the writer, however, the questions that may be answered amply justify the labor as well as the necessity for delaying publication.

REFERENCES

1. Bain, R.: "Religious attitudes of college students." *Amer. J. Sociol.*, Vol. XXXII, 1927, pp. 762–770.
2. Boldt, W. J., and Stroud, J. B.: "Changes in the attitudes of college students." *J. Ed. Psychol.*, Vol. XXV, 1934, pp. 611–619.

3. Carlson, H. B.: "Attitudes of undergraduate students." *J. Soc. Psychol.*, Vol. V, 1934, pp. 202–213.
4. Cowley, W. H.: *The Personnel Bibliographical Index.* Columbus, Ohio, Bureau of Educational Research, Ohio State University, 1932, 433 pp.
5. Droba, D. D.: "Education and Negro Attitudes." *Sociol. & Soc. Res.*, Vol. XVII, 1932, pp. 137–141.
6. Droba, D. D.: "Effect of various factors on militarism-pacifism." *J. Abn. & Soc. Psychol.*, Vol. XXVI, 1931–32, pp. 141–153.
7. Droba, D. D.: "A scale of militarism-pacifism." *J. Educ. Psychol.*, Vol. XXII, 1931, pp. 96–111.
8. Dudycha, G. J.: "The beliefs of college students concerning evolution." *J. App. Psychol.*, Vol. XVIII, 1934, pp. 85–96.
9. Edgerton, H. A., and Toops, H. A.: *Academic Progress.* Columbus, Ohio, Ohio State University Press, 1929, 150 pp.
10. Garrison, K. C., and Burch, J. S.: "A study of racial attitudes of college students." *J. Soc. Psychol.*, Vol. IV, 1933, pp. 230–235.
11. Garrison, K. C., and Mann, Margaret: "An analytic study of the opinions of college students." *J. Soc. Psychol.*, Vol. II, 1931, pp. 168–178.
12. Harper, Manly H.: *Social beliefs and attitudes of American educators.* Teachers College, Columbia University, Contributions to Education, No. 294, 1927, 91 pp.
13. Harris, A. J., Remmers, H. H., and Ellison, C. E.: "The relation between liberal and conservative attitudes in college students, and other factors." *J. Soc. Psychol.*, Vol. III, 1932, pp. 320–336.
14. Johnson, C. S.: "Racial attitudes of college students." *Publ. Amer. Sociol. Soc.*, Vol. XXVIII, 1934, pp. 24–31.
15. Jones, E. S.: "The opinions of college students." *J. App. Psychol.*, Vol. X, 1926, pp. 427–436.
16. Katz, Daniel, and Allport, F. H.: *Students' Attitudes.* Syracuse, N. Y., Craftsman Press, Inc., 1931, 408 pp.
17. Kellar, Beatrix: "The construction and validation of a scale for measuring attitude toward any home-making activity." *Studies in Higher Education XXVI, Bulletin of Purdue University*, Vol. XXXV, No. 4, 1934, pp. 47–63.
18. Kolstad, A.: *A study of opinions on some international problems.* Teachers College, Columbia University, Contributions to Education, No. 555, 1933.
19. Salner, E., and Remmers, H. H.: "Affective selectivity and liberalizing influence of college courses." *J. App. Psychol.*, Vol. XVII, 1933, pp. 349–354.
20. Sowards, Genevieve S.: "A study of the war attitudes of college students." *J. Abn. and Soc. Psychol.*, Vol. XXIX, 1934–35, pp. 328–333.
21. Symington, T. A.: *Religious liberals and conservatives.* Teachers College, Columbia University, Contributions to Education, No. 640, 104 pp.
22. Symonds, P. M.: "A social attitudes questionnaire." *J. Educ. Psychol.*, Vol. XVI, 1925, pp. 316–322.
23. Thurstone, L. L., and Chave, E. J.: *The Measurement of Attitude.* Univ. of Chicago Press, 1929, 97 pp.
24. Vetter, G.: "The measurement of social and political attitudes and the related personality factors." *J. Abn. and Soc. Psychol.*, Vol. XXV, 1930, pp. 149–189.
25. Willoughby, R. R.: "A sampling of student opinion." *J. Soc. Psychol.*, Vol. I, 1930, pp. 164–169.

STUDIES OF JEWISH AND NON-JEWISH INTELLIGENCE

MOSHE BRILL

Ohio State University

The problem of the intelligence of Jews as compared with that of non-Jews has ever been one of great interest and controversy. There are contradictory opinions in the literature regarding the differences between Jews and non-Jews in (1) average general mental level, (2) variability of the distributions, (3) frequency of mentally superior and inferior, (4) growth of general intelligence, and (5) manual ability.

Two summaries[1] of the results of comparative studies of the intelligence of Jews and non-Jews have been reported. Both of these expressed disapproval of the sweeping generalizations made by investigators, but did not present a critical evaluation of the studies reported.

The present study is a critical resumé of the literature, primarily experimental, on the intelligence of Jews as compared with that of non-Jews. The procedure here followed was to evaluate critically the assumptions, method and procedure, and statistical evidence of each investigation reported. A general summary of the conclusions arrived at through the critical resumé of all the studies reported will follow.

The research literature on the intelligence of Jews and non-Jews may be classified into (1) studies in which no specific comparison of Jews and non-Jews was made, but which refer to the relative mental level of the two groups, and (2) comparative studies of the performance of Jews and non-Jews on standardized intelligence tests. Practically all of the studies reported in England and the United States will be presented in the order of their publication. The English studies will be presented first.

STUDIES IN ENGLAND

The Study by Pearson and Moul.[2]—Pearson and Moul reported a survey in which Jewish children in a London school were rated by their teachers by means of the Pearson Biometric Scale. Comparing these distributions with those of the London Pioneer Survey, the authors concluded: "The Jewish girls have less intelligence than the Gentile girls in any type of council school. The comparisons of Gentile

and Jewish boys is less clear-cut . . . they (the Jews) are not as good as the boys of the medium or average school, but are better than the boys of the poor type of school."

TABLE I.—THE DISTRIBUTION OF INTELLIGENCE OF JEWISH AND NON-JEWISH CHILDREN IN LONDON SCHOOLS AS MEASURED BY THE BIOMETRIC SCALE

Source of data	Groups	Boys					Girls				
		N	*V*	*I*	*S*	*D*	*N*	*V*	*I*	*S*	*D*
Pearson and Moul	Jews	619	16.0	36.2	34.7	13.1	585	7.5	25.5	37.8	29.2
L.P.S.* poor schools	Non-Jews	...	15.6	32.3	23.5	28.6	...	15.8	37.2	29.4	17.6
L.P.S. medium schools	Non-Jews	...	21.3	37.2	24.7	16.8	...	12.4	41.4	32.5	13.7
L.P.S. better schools	Non-Jews	...	18.6	34.8	23.1	23.5	...	13.9	34.6	25.0	26.5
Davies and Hughes	Jews	232	21.5	42.7	25.4	10.4	251	13.9	32.3	23.5	30.3
School *A*	Non-Jews	166	7.8	34.3	35.0	22.9	164	9.1	25.0	31.1	34.8
Davies and Hughes	Jews	155	44.2	35.5	11.6	7.7	171	26.9	31.6	20.5	21.0
School *B*	Non-Jews	115	13.9	36.5	23.5	26.2	114	7.0	26.3	24.6	42.2
Davies and Hughes	Jews	148	16.9	49.3	26.4	7.4	122	15.5	48.4	25.4	10.7
School *C*	Non-Jews	115	5.2	38.3	33.0	23.5	134	3.7	20.1	27.6	48.5

Legend: *V*, very able, capable; *I*, intelligent; *S*, slow; *D*, dull, very dull.

* L.P.S., London Pioneer Survey. The distribution of the mental levels of the girls in the poor schools of the L.P.S. is peculiar, as the mental calibre of these girls appears to be superior to that of the girls in medium and better schools.

Pearson did not report the differences between the groups in variability of mental ratings. The combined per cent frequencies in the categories *I* and *S* for Pearson's Jewish boys and girls were 70.9 and 63.3, respectively, while those of the non-Jewish boys and girls in the L.P.S. better schools were 57.9 and 59.6, respectively, that is, the distributions of the Jewish children were more compact.

The main criticisms of this study are: (1) The ratings were made by different teachers. (2) Teachers' ratings of the intelligence of foreign children are notoriously unreliable. (3) The Jewish and non-Jewish children were not of the same socio-economic status. Only 6 per cent of the fathers of the Jewish children, for instance, were native-born or naturalized, and 71.9 per cent of them were reported to be engaged in occupations of low-grade labor. (4) Pearson's contention that the father's occupation is no criterion of the innate intelligence of the child is unfounded. (5) The reliability and validity of the Biometric Scale for measurement of intelligence has not been established.

When one compares the Jewish children with the non-Jewish children in the poor schools of the L.P.S., the following conclusions may be drawn: (1) The average mental level of the Jewish boys is negligibly higher than that of the non-Jewish boys. In the case of the girls, the reverse is true. (2) The distributions of the intelligence levels of the Jewish children are more compact than those of the non-Jewish children. (3) There are as many Jewish as non-Jewish boys rated *Very able and capable*, but less than one half as many Jewish as non-Jewish boys rated *Dull and very dull.* In the case of the girls, half as many Jewish as non-Jewish are rated *Very able and capable*, and more of the Jewish than the non-Jewish girls rated *Dull and very dull.*

The Study by Davies and Hughes.[3]—Davies and Hughes compared Jewish and non-Jewish children, ages eight to fourteen, in intelligence and achievement in English and arithmetic, using the Northumberland Standardized Tests (1925 Series). These schools in London provided, in the estimation of the authors, the best approximation to a similarity of social and economic status in Jewish and non-Jewish families. School *A* was situated in a good district, School *B* in a moderately poor district, and School *C* had the poorest group of Jewish children in London.

The authors concluded: "The most striking feature is the general superiority of the Jewish children, whether boys or girls." Appreciable differences in favor of the Jewish children were found in ninety-three groups out of one hundred compared. Assuming as their criterion of significance that the difference must be at least three times its *PE*, the authors found that fifty-two per cent of the differences were significant.

The Jewish boys were distinctly in advance of the non-Jewish boys in the same school, the differences being relatively constant in the three schools. At age ten, the difference was equivalent to over one year. The differences between the girls were considerably less in the good district than in the poorest district. There was a decided drop at age twelve in the differences between the Jewish and non-Jewish children. The authors do not contribute this drop to "a precocious arrest of development, for the Jewish superiority is more than regained by the boys of thirteen, and a similar tendency, although not so marked, is shown by the girls." More Jewish than non-Jewish children about the age of eleven are transferred to secondary schools. It may also be due, they suggest, to the fact that pubertal influences

TABLE II.—A SUMMARY OF DAVIES AND HUGHES' DATA CONCERNING THE INTELLIGENCE OF THE JEWISH AND NON-JEWISH CHILDREN

	School *A*				School *B*				School *C*			
	Jews		Non-Jews		Jews		Non-Jews		Jews		Non-Jews	
	Range	Median	Range	Median	Range	Median	Range	Median	Range	Median	Range	Median
Boys												
Average mark	134–256	216	84–251	199	143–252	230	115–221	175	99–232	190	52–215	150
SD	33–56	49	41–63	55	27–65	46	30–69	61	44–60	53	39–74	61
N	7–46	37	4–37	24	6–38	23	7–27	16	2–33	24	2–25	19
Girls												
Average mark	116–234	224	114–235	209	108–255	216	66–266	189	104–228	172	56–182	142
SD	33–55	47	39–60	55	45–71	55	30–60	50	16–64	54	45–64	53
N	4–52	37	1–30	26	5–36	26	1–30	18	5–25	18	5–26	22

occurred at different ages. The Jewish boys were slightly superior to Jewish girls; this result was reversed among non-Jewish children. The teachers' estimates of the intelligence of the same children, using the Biometric Scale, corroborated the results of the standardized tests. In computing the reliability of the differences between the sigmas of the distributions of the Jewish and non-Jewish children, the present writer found that out of thirty-six differences, twenty-one were larger in the case of the non-Jews, two of them were reliable, and four others approached significance; fifteen of the differences were larger in the case of the Jews, and of these only two approached significance.

In order to rule out the factor of socio-economic status, Hughes[4] compared the Jewish and non-Jewish children of the previous study in groups arranged according to the following types of occupational status of the parents:

1. Manufacturers, merchants, travellers, rabbis, teachers, higher grade clerks.
2. Engineers, electricians, cabinet-makers, metal workers.
3. Shopkeepers, dealers, and small masters.
4. Policemen, postmen, bus and taxi drivers, conductors, soldiers, sailors, railway porters.
5. Makers of boats, clothes, umbrellas, buttons, boxes, cigarettes.
6. Bricklayers, carpenters, painters, plumbers.
7. Shop assistants, hairdressers, tobacconists.
8. Carmen, coalmen, caretakers.
9. Laborers (mainly in docks), hawkers.
10. No father.
11. No occupation.

TABLE III.—SUMMARY OF THE STATISTICAL DATA AS ANALYZED BY HUGHES

	Occupational groups										
	1	2	3	4	5	6	7	8	9	10	11
Jews.....	109.7	107.7	110.1	105.4	103.9	105.1	102.0	112.4	103.8	100.6	85.8
Non-Jews....	103.5	101.3	98.4	101.5	97.3	93.7	88.7	92.8	90.1	93.0	90.7

In groups 4, 6, 8, 9, and 11, the numbers of Jewish children were too small to give any reliable results. A steady superiority in the mental ratios of the Jewish children was found in groups 1, 2, 5, and 10, amounting to more than half a year at age ten. In groups 3 and 7, the difference was about twice this amount. Hughes feels that the

result, especially in group 3, seems to indicate that "commerce attracts Jews of good intelligence, men of the calibre, who, if they were non-Jews would probably enter the professions." He concluded that the poorest Jewish children were practically equal to the general non-Jewish average, "a fact which points very definitely to 'racial' superiority." He also concluded that even after allowance has been made for the differences in socio-economic status, "there still remains a 'racial' difference equivalent at age ten to about 0.7 mental years."

The main criticisms of the studies by Davies and Hughes are: (1) The samples in most age groups were too small to permit generalizations. (2) The factor of language handicap was not ruled out.

The following conclusions may be drawn from these studies: (1) The Jewish children were superior to the non-Jewish children of the same age, school, and similar socio-economic status, the superiority being more marked with the boys than with the girls. (2) The distributions of the Jewish children were on the whole more homogeneous than those of the non-Jewish children. (3) The Jewish children had larger frequencies than the non-Jewish children in the category *Very able and capable*, and smaller frequencies in the category *Dull and very dull.*

The Study by Winch.[5]—Winch compared Jewish and non-Jewish children, ages six to fourteen, in East End London on the Winch Reasoning tests, sets *A* and *E*, and a five-point handwriting scale as a measure of dexterity. The children were rated on a five-point "Social Class" scale, according to the aggregate of their social class conditions, such as family rental and occupation of the parents.

The author's main conclusions are: "The differences in intelligence between Jews and Christians are greater in the male than in the female sex." "Jewish boys and girls of elementary-school age are more advanced for their ages and more intelligent, but so far as handwriting is indicative of manual dexterity, the Jewish boys appear about the same and the Jewish girls definitely inferior to the Christian girls." "The variability of the Jews is slightly less than that of the Christians in writing on the whole, as far as the boys are concerned, and more in writing as far as the girls are concerned."

The main criticisms of this study are: (1) The samples of both groups were very small. (2) A larger proportion of Jews than non-Jews at age 11+ obtained scholarships, attended central schools, or had become fee-paying pupils at secondary schools. The author did not report the effect of this "important issue" on the samples. (3)

TABLE IV.—A SUMMARY OF WINCH'S DATA

	Median average marks								Median coefficient of variability		
	N		*Rs*		*Wr*		*SC*				
	R	*M*	*R*	*M*	*R*	*M*	*R*	*M*	*Rs*	*Wr*	*SC*
Jewish boys	2–28	18	.32–3.00	2.57	.75–3.00	2.25	2.33–3.50	2.71	.39	.30	.42
Non-Jewish boys	1–24	15	.13–3.17	1.59	1.00–2.84	2.50	1.00–2.07	2.00	.56	.35	.58
Jewish girls	2–31	20	.31–3.00	2.13	1.92–2.88	2.58	1.78–2.74	2.40	.43	.25	.42
Non-Jewish girls	4–24	18	.13–2.79	1.61	1.86–2.88	2.59	1.13–2.04	1.55	.45	.23	.56

Legend: *Rs*, reasoning; *Wr*, writing; *SC*, social class; *R*, range; *M*, median.

The validity and reliability of the ratings of "Social Class" and of the Handwriting test were not reported. (4) The Jewish children had a higher Social Class rating than the non-Jewish children, but the author did not report how this factor affected the difference in intelligence between the groups. (5) No evidence is reported for the author's assumption that handwriting is an adequate means for the measurement of manual dexterity.

The following conclusions may be drawn from this study: (1) The Jewish children were superior to the non-Jewish children in scores on the Reasoning tests, the difference between the boys being greater than that between the girls. In handwriting, Jewish and non-Jewish boys were more nearly alike, and non-Jewish girls superior to Jewish girls. (2) The variability of the distributions of the Jewish children was smaller than that of the non-Jewish children, except in the case of the writing of the girls.

STUDIES IN THE UNITED STATES

1. *Studies in Which No Specific Comparisons of Jews and Non-Jews Were Made, but Which Refer to the Relative Intelligence of Jews and Non-Jews.*—Such studies as the ones by Weintrob and Weintrob,[6] Van Denberg,[7] Judd,[8] Brigham,[9] Halpern,[10] and Maller[11] are often quoted as evidence regarding the native ability of Jews. These studies are at best indirect evidence of questionable value. This type of investigation will be illustrated by Brigham's[9] study of Jewish intelligence.

Brigham, in his study of the intelligence of army recruits, had no separate distribution of the scores of the Jews. In view of the fact

that in the 1910 census about fifty per cent of those reporting Russia as their country of origin spoke Hebrew or Yiddish, Brigham stated:

> Our figures, then, would rather tend to disprove the popular belief that the Jew is highly intelligent. Immigrants examined in the army, who report their birthplace as Russia, had an average intelligence below those from all other countries except Poland and Italy. . . . The sample from Russia had a higher standard deviation than that of any other immigrant group sampled. . . . If we assume that the Jewish immigrants have a low average intelligence, but higher variability than other native groups, this would reconcile our figures with popular belief, and at the same time, with the fact that investigators searching for talent in New York City and California schools, find a frequent occurrence of talent among Jewish children. The able Jew is probably recognized, not only because of his ability, but because he is able and a Jew.

The round-about way taken by Brigham to determine the relative intelligence level of Jews is indeed open to serious objections and criticisms. Hirsch[12] found that Russian Jews scored reliably higher than Russian non-Jews in this country. Even assuming that fifty per cent of the recruits reporting Russia as their country of origin were Jewish, it stands to reason that the non-Jews in this group pulled the average down, whereas the Jews with their high scores caused an increase in the variability of the distribution. Since Brigham included Polish, Russian, Lithuanian, and other Jews in his Russian group, it is doubtful whether his conclusions regarding 'Russian' intelligence are valid. He made no allowance for the language handicap of the 'Russian' group. He assumed that the more recent immigrants were of innately lower intelligence. By reason of the greater restrictions on immigration year by year, it seems more plausible to assume the opposite to be true. Brigham later admitted that this study "with its entire hypothetical superstructure of racial differences collapses completely."[13]

2. *Comparative Studies of the Intelligence of Jews and Non-Jews. The Study by Grier.*[14]—Grier compared Jewish and non-Jewish high-school students, and reported that when grouped by ages, a noticeable superiority in school standing "as shown by the curriculum test was found to be with the Gentiles." This superiority was observed to appear most strongly at the earlier ages, slowly decreasing upwards, but at the age of sixteen, Jewish boys were slightly superior. The Jewish girls were reported to be mentally superior at most ages, but the non-Jewish girls showed a higher scholarship at the age of fourteen.

He did not report the statistical or any other data upon which he based his conclusions.

The Study by Murdock.[15]—Murdock compared Jewish, Italian, native white, and Negro boys on the Pressey Group Intelligence Test. At every age except thirteen and fourteen, the Jewish group scored negligibly higher than the native white group. The fifty percentile of the Italians at no age equalled the twenty-five percentile of the Jews, and the fifty percentile of the Jews at each age exceeded the seventy-five percentile of the Italians. The differences in variability between the Jews and native white were unreliable.

TABLE V.—A SUMMARY OF MURDOCK'S DATA

Age.....	9	10	11	12	13	14	15	9	10	11	12	13	14	15
	Median								Semi-interquartile-range*					
Jews.....	109.0	109.3	118.4	128.5	125.5	124.3	126.5	11.4	15.7	13.0	13.8	13.5	15.0	7.5
Italian.....	73.5	84.3	94.8	105.5	109.5	109.5	113.5	12.7	11.5	15.3	16.5	14.5	14.7	18.5
Native.....	108.5	108.5	111.0	127.0	131.0	128.5	120.0	....	6.3	20.0	15.3	12.3	14.5	13.0

* Computed by the present writer from Murdock's data.

The main criticisms of this study are: (1) The groups compared were not equated in socio-economic status. The Italian group was selected. (2) No evidence is given as to whether or not the factor of language handicap was ruled out. (3) The groups compared were not adequate samples of their respective national groups.

The following conclusions may be drawn from this study: (1) The Jewish and the native white groups scored fairly equal, and both groups exceeded the Negroes and Italians, especially the latter. (2) The differences in variability between the Jewish and native white groups were unreliable.

The Study by Jordan.[16]—Jordan compared elementary-school children of Roumanian Jewish, Russian Jewish, and native parentage in teachers' marks, amount of retardation and acceleration, and in scores on the Trabue Language Scale *B* and *C*, opposites, vocabulary, substitution, memory span, geometric forms, and number completion tests.

The school marks, Jordan stated, "while appearing conclusive, are really masking the real situation." She concluded: "There are no great national differences apparent in the abilities which do not

TABLE VI.—A SUMMARY OF JORDAN'S DATA

	Percentage of			Scores in all tests				
				Language		Non-language		
	N	Acceler-ation	Retar-dation	Medi-an	SD	Medi-an	SD	*N*
Native white	350	38.8	32.0	271.4	47.5	597.5	122.0	129
Roumanian Jews	99	51.5	24.2	242.5	47.5	565.0	140.2	54
Russian Jews	336	30.3	45.5	242.5	51.5	607.5	138.0	81

require language knowledge or familiarity. The only exception to this is in the case of perception of geometric forms, in which evidently the racial trend of the Jewish people away from mechanics and mechanical ingenuity seems to be productive of a definite loss in form-perception."

The main criticisms of this study are: (1) The groups compared were not equated in socio-economic status. The percentage of fathers engaged in the professions for the native white, Roumanian Jewish, and Russian Jewish groups were 18.1, 0.0, and 2.8 respectively. (2) The factor of language handicap was not ruled out. (3) It is doubtful whether the tests used were really measures of intelligence or mechanical ability. (4) The author took the liberty of shortening the time allowance for all but the Trabue test.

The following conclusions may be drawn from this study: (1) The total score of the native white on the language tests was higher than that of the Jewish groups, while the Russian Jewish group scored higher than the native white on the non-language tests. The Roumanian Jewish group ranked highest in amount of acceleration and lowest in retardation. (2) The Roumanian Jewish group varied more in total scores on the language tests; the Russian Jews had a larger variability than the other groups in total scores on the non-language tests.

The Study by Berry.[17]—In classifying *IB* pupils by means of the Detroit Primary Intelligence Examination, Berry found that "twice as large a per cent of the pupils who come from homes where Yiddish is spoken is found in the *X* group, and less than one-half as large a per cent in the *Z* group as of the pupils who come from homes where Italian is spoken." "Of the groups coming from homes where

a foreign language is spoken, the Jews not only test the highest, but have the largest per cent completing a semester's work and the largest per cent of promotion."

TABLE VII.—INTELLIGENCE IN RELATION TO HOME LANGUAGE

	Frequency per cent in groups			*N*	Per cent of total number
	X	*Y*	*Z*		
English white	30.3	59.6	10.4	6186	60.0
Yiddish	13.2	68.9	17.9	622	6.0

The main criticisms of this study are: (1) Nothing is reported concerning the Jewish children who gave their home language as other than Yiddish. It is doubtful, therefore, whether the Yiddish-speaking group was an adequate sample of the entire Jewish school population studied. The lower mental groups of the foreign-born population tend, as a rule, to use their native tongue even after a long period of residence in the new country. (2) Nothing is reported concerning the socio-economic status of the language groups compared.

The following conclusions may be drawn from this study: (1) The Yiddish-speaking group scored lower than the native white, but higher than other foreign-speaking groups. (2) The distribution of the scores of the Yiddish-speaking group was more compact.

The Study by Bailey.[18]—The per cent frequencies of mentally deficient among the volunteers and draftees to the last World War for the United States average and for the Hebrew groups were 29.2 and 17.2 respectively. Next to the Scotch and Welsh, the Hebrew group showed the lowest percentage of mental defect.

The Study by Pintner and Keller.[19]—The mean Binet IQ's for groups of kindergarten, first- and second-grade American white, English-speaking, and foreign-speaking children were ninety-five, ninety-two, and eighty-four respectively. The average for all the Jewish children speaking English or a foreign tongue was 95.0. Nothing is reported regarding the relative socio-economic status and knowledge of English of the groups compared. It is highly doubtful whether the Jewish group in this study was representative.

The Study by Dearborn.[20]—In a community studied by Dearborn, he found that "The Jewish children rank highest in mental age and in intelligence quotients, then come the American children, and last

the Italians. In anatomical development . . . the Jewish children are far ahead of the other two racial groups." He reported no statistical data concerning the groups compared.

The Study by Bere.[21]—Bere compared samples of one hundred ten-year old Italian, Hebrew, and Bohemian boys in scores on the Stanford Binet, a short Pintner-Patterson, NIT, and the Pintner Non-Language Mental Test. The elimination of intra-group variations as to social status and language used in the home did not eliminate the differences between the groups, neither did it change their relative rank positions. Bere found a zero correlation between the Binet and the Pintner-Patterson scores and average length of residence in the United States. She also found a rise in the Binet and performance-test ages with increase use of English in the home. This, she believes, indicates definitely that a language handicap is not responsible for low Binet IQ's. The correlations between the individual tests and the occupational status of the parents were all found to be positive, and varying in size for the different groups. It is difficult to interpret these findings "because of the inadequacies of the scales and limitations set by getting this information from the child."

TABLE VIII.—A SUMMARY OF THE STATISTICAL FINDINGS BY BERE

Test	Mean			Median			SD			*Q*			*N*			D/PED*		
	H	*I*	*B*	*H*	*I*	*B*	*H*	*I*	*B*	*H*	*I*	*B*	*H*	*I*	*B*	*H & I*	*H & B*	*I & B*
1	123	108	118	119	105	116	16	12	15	9.0	6.5	8.0	100	100	100	15.99	5.55	10.60
2	98	85	93	95	83	92	13	10	10	8.0	5.0	6.5	100	100	100			
3	102	109	123	98	105	116	20	21	16	11.5	13.5	19.0	100	100	100	4.91	27.30	11.02
4	79	46	...	78	46	...	25	21	..	18.0	17.0	....	246	272	...	27.26		
5	298	264	...	297	263	...	70	78	..	55.0	53.0	....	246	272	...	9.21		
6	...	...	...	102	81	...	..	..	..	11.0	9.0	....	246	272	...	29.90		
7	...	...	...	58	52	...	..	..	..	8.5	8.0	....	...	...	...	10.47		

Key to tests: 1, Binet TA; 2, Binet IQ; 3, Pintner-Patterson TA; 4, NIT score; 5, Pintner non-language score; 6, NIT IQ; 7, Pintner non-language index.

* In tests 1 and 3 the *D* is between the means; in others it is between the medians.

The main criticisms of this study are: (1) It is doubtful whether her samples of Jewish and Italian children were random and representative. (2) The Jewish and Italian groups differed considerably in socio-economic status.

The following conclusions may be drawn from this study: (1) The three groups ranked in the following order on the Binet test: Hebrew,

Bohemian, and Italian; in the Pintner-Patterson scale they ranked: Bohemian, Italian, and Hebrew. In the NIT and in the Pintner non-language test the Hebrews scored reliably above the Italians. The distributions of the Hebrews were more heterogeneous than those of the other groups in the Binet and NIT; in the Pintner-Patterson scale and in the Pintner non-language test the distributions of the Hebrews were more homogeneous than those of the Italians.

The Study by Feingold.[22]—Feingold compared high-school students of native, Jewish, and seven other national groups in IQ's on a modified Army Alpha.

TABLE IX.—FEINGOLD'S DATA ON NATIVE AND JEWISH INTELLIGENCE

	Freshmen		Juniors		Seniors	
	N	IQ	*N*	IQ	*N*	IQ
Native white	892	103	286	107	264	100
Jewish	518	103	208	103	146	96

Feingold concluded: "The mental differences among the American reared descendants of foreign races . . . are so small as to be practically insignificant."

The main criticisms of this study are: (1) Nothing is reported concerning the reliability and validity of the test used. One does not know whether a difference of a few points is at all significant. (2) In computing the IQ's, Feingold divided the MA by the CA, the latter being any number above even eighteen years. His method in this respect is not clearly stated, but may be deduced by checking some of his IQ's. (3) Nothing is reported concerning the socio-economic status of the groups compared.

The Study by Fukuda.[23]—Fukuda found the mean Binet IQ of his Jewish group to be 87, but this sample consisted of *only three children.*

The Study by Terman.[24]—10.5 per cent of Terman's main experimental group of gifted children were Jewish. He stated: "The actual per cent is probably greater than this, as there is reason to believe that the presence of Jewish blood has in some cases been concealed . . . the amount of Jewish blood in our groups is about twice the expected."

The Study by Luckey.[25]—Luckey reported the distributions of Terman IQ's for native white, Jewish, and eight other racio-national

groups, all of whom were clinic cases. The main criticisms of this study are: (1) It is highly doubtful whether the clinic population may be considered an adequate sample of the entire population. (2) The groups compared were not equated in anything, except, perhaps, in the fact that all of them were clinic cases.

The following conclusions may be drawn from this study: (1) The Jewish and native white scored very nearly alike. (2) There was no appreciable difference between them in variability. (3) Negligible differences were found between them in frequencies of high and low scores.

TABLE X.—LUCKEY'S DATA CONCERNING NATIVE WHITE AND JEWISH INTELLIGENCE

	Per cent frequency of IQ levels								*N*
	25–50	50–65	65–80	80–90	90–110	110–120	120–140	Over 140	
Native white....	1.38	7.56	21.56	19.12	28.71	11.84	8.81	1.01	2964
Jewish.....	1.96	9.08	17.79	18.28	28.71	12.27	10.80	1.10	815

The Study by Seago and Koldin.[26]—Twelve-year old Jewish and Italian children were compared in scores on the NIT, and in scores on the five individual tests of the NIT series. The authors concluded that the superiority of the Jewish over the Italian group "corresponds to that found for young bright over old dull children of American or British birth, which suggest that the Italians as a group represent a lower level of development than the Jews, or that the difference is one of general mental level rather than of innate special abilities." The authors reported no evidence to the effect that the two groups were really representative samples of their respective national groups, neither did they report how the two groups compared in socio-economic status.

The following conclusions may be drawn from this study: (1) The Jewish children scored reliably higher than the Italian. (2) The difference between the two groups in variability was insignificant.

The Study by Graham.[27]—Graham compared white and Jewish children, Habit Clinic cases, in scores on the Stanford Binet, Porteus Maze, Knox Cube, Healy PCI, Mare and Foal, Diagonal Formboard, Healy Construction *A*, and Mannikin tests. He concluded: "Whether

TABLE XI.—A SUMMARY OF SEAGO AND KOLDIN'S DATA

	Factors kept constant												
	Age		Age, grade, sex							Age, grade, sex, NIT total			
N.....	J-800, I-452		J-250, I-143							J-100, I-100			
Tests.....	1		1		2	3	4	5	6	2	3	4	5
	Mean	SD	Mean	SD	Mean								
Jews.....	105.7	221.1	102.5	20.5	16.0	23.2	25.9	16.6	20.9	15.2	21.8	24.2	14.7
Italians....	88.5	22.3	90.4	20.8	14.8	20.0	23.1	12.2	20.4	15.4	20.9	24.2	13.7
D/PED....	19.4	.....	8.3	....	4.6	7.3	5.4	8.0	1.1	.8	1.7	0	1.4

Tests: 1, NIT total score; 2, arithmetical reasoning; 3, sentence completion; 4, logical selection; 5, synonym-antonym; 6, symbol digit.

measured by verbal or non-verbal tests, the Jewish are found to be superior—superior even to the American clientele of the Habit Clinics." The results of these data, he believes, are not to be taken as of a special group, in view of the fact that the IQ's are on a par with those of the unselected population. One may doubt this contention. Nothing is reported regarding the socio-economic status of the groups compared.

TABLE XII.—GRAHAM'S DATA CONCERNING NATIVE WHITE AND JEWISH INTELLIGENCE

	CA in months			Binet					Porteus		Knox-Cubes scores		Healy PCI	
				TA in months			IQ							
	M	SD	*N*	*M*	SD	*N*	*M*	SD	Six years	Nine years	Six years	Nine years	Six years	Nine years
Native.....	76	37	60	74	40	59	99	24	5.9	7.7	3.4	4.8	135	235
Jews.....	69	30	47	72	33	47	105	23	5.9	8.1	3.8	5.6	206	278

The following conclusions may be drawn from this study: (1) The Jewish children scored higher on both the Binet and non-verbal

tests than the native white; even when the factors of chronological and Binet ages were partialled out, the relative position of the Jewish children remained unchanged. (2) The Jews were slightly more homogeneous than the natives in Binet TA's and IQ's.

The Study by Hirsch.[12]—Hirsch compared native white, Polish Jewish, and Russian Jewish children in IQ's on the Pintner-Cunningham Primary Mental Test (P-C), and the revised Dearborn Group Tests of Intelligence, Series I, General Examination *A* (D-2) and *B* (D-3). The Jewish groups scored higher than the native white. The difference between the native white and Polish Jews was almost conclusively reliable. The differences between the Jewish and native white were primarily in D-3 given to the fourth to ninth grades. In percentage frequency at or above one hundred twenty IQ the groups ranked: Polish Jews, Russian Jews, and natives; when the percentage of the groups falling below seventy-nine IQ was compared the rank was reversed. According to Hirsch the differences in intelligence were national or natio-social rather than social. Rumyaneck[28] questioned this belief, and stated: "It is difficult to see, if this is the case, why Polish Jews rank higher than Russian Jews. Racially, both groups are very much alike." However, this difficulty is not a serious one inasmuch as the differences between the averages of the Polish and Russian Jews are unreliable.

TABLE XIII.—A SUMMARY OF HIRSCH'S DATA CONCERNING THE INTELLIGENCE OF THE NATIVE WHITE, POLISH JEWISH, AND RUSSIAN JEWISH GROUPS

	IQ		*N*	D/PED		IQ in tests			*X*	*Y*	*Z*
	Mean	SD		Polish Jews	Russian Jews	P-C	D-2	D-3			
Native.....	98.3	15.87	1030	3.6	2.3	97.8	97.3	98.8	8.4	11.1	
Polish Jews.....	102.8	14.55	75	...	2.6	....	99.6	104.0	13.3	8.0	70.7
Russian Jews....	99.5	14.58	627	2.6	...	97.7	97.9	100.2	9.1	9.9	52.0

Legend: *X*, per cent at or above one hundred twenty IQ; *Y*, at or below seventy-nine IQ; *Z*, at or above native average.

The main criticisms of this study is that Hirsch did not control the factors of socio-economic status and knowledge of English of the groups.

The following conclusions may be drawn from this study: (1) The Jews had a higher IQ than the natives. (2) The natives tended to greater heterogeneity than the Jews. (3) The Jews had a larger frequency of high scores and a smaller frequency of low scores than the native white.

The Study by Goodenough.[29]—Goodenough studied the Draw A Man IQ's of children of other racial or national groups. Of all the racial groups compared the Jewish ranked highest in IQ and lowest in coefficient of variability.

TABLE XIV.—THE GOODENOUGH IQ'S OF NATIVE WHITE AND JEWISH CHILDREN

	Mean IQ	*N*	SD	Coefficient of variability
Native white	101.5	500	18.3	18.0
Jews	106.1	55	16.2	15.3

The main criticisms of this study are: (1) The groups were not compared in socio-economic and cultural status. (2) The Jewish groups was not, as Goodenough herself indicated, an adequate sample.

The Study by Cohen.[30]—Jewish and non-Jewish freshmen at Ohio State University were compared in total scores on the OSU Psychological Examination.

TABLE XV.—SUMMARY OF COHEN'S DATA

	Jewish men	Jewish freshmen	Non-Jewish men	Non-Jewish freshmen	Jewish men and freshmen	Non-Jewish men and freshmen	Jewish and Non-Jewish men	Jewish and Non-Jewish freshmen
Mean	51.9	59.6	47.1	45.4				
SD	24.1	22.9	27.9	26.7				
N	128	65	135	58				
D/PE_D	.....	....	.....	....	3.2	5.9	2.6	8.3

The order of superiority in both total scores and scores on the sub-tests was: Jewish women, Jewish men, non-Jewish men, and non-Jewish women. In sixteen out of eighteen comparisons made on the sub-tests, the Jewish groups were found superior; in ten of these, the

differences were reliable, and in three they approached reliability. Cohen concluded: "The results seem to indicate a definite and dependable Jewish superiority, especially on the tests which necessitate language ability."

The main criticisms of this study are: (1) The groups were not equated in socio-economic status. (2) It is doubtful whether Jewish college students are as representative of their national group as are non-Jewish college students of theirs.

The following conclusions may be drawn from this study: (1) The Jewish men were found superior to the non-Jewish men, but this difference was not reliable. The Jewish women were found superior to the non-Jewish women, the difference being reliable. (2) The non-Jews were more variable than the Jews.

The Study by Hayes.[31]—Of four hundred thirty-five students at the University of Illinois, Hayes found that the Jewish group (N-93) had an average scholastic grade of 3.18 as compared with that of the Chinese (N-36) 3.35, and the Nordics (?) (N-100) 3.00. The average scholastic grade for the entire group was 2.93. Two of the ten students with the highest grade were Jewish. The author did not use standardized intelligence or achievement tests. He considered a difference of .18 in scholastic grades, for instance, large enough to warrant conclusions.

The Study by Rigg.[32]—Rigg compared children in grades III to VIII, of native white and Jewish parentage in scores on the NIT, Woody-McCall Arithmetic Fundamentals, and Thorndike McCall Reading Test. The groups were not equated in socio-economic status or in knowledge of English.

TABLE XVI.—RIGG'S DATA CONCERNING NATIVE WHITE AND JEWISH SCORES

	N	Median quotients ± PE Md		
		Intelligence	Arithmetic	Reading
Native	8130	104.85 ± .17	102.96 ± .15	100.28 ± .17
Jewish	445	103.19 ± .82	106.10 ± .72	98.35 ± .76
D/PED	....	1.98	4.24	2.47

The following conclusions may be drawn from this study: (1) The Jewish children scored insignificantly lower than the native children in the NIT and reading test; they scored reliably higher on the arithmetic test.

The Study by Garrett.[33]—Garrett compared Jewish and native white Columbia College freshmen in grades, Thorndike Intelligence Examination, George Washington Social Intelligence, and the Laird Personal Inventory B2. He concluded: "The Hebrew students rank highest in grades and Thorndike Examination. Here they rank highest too in Social Intelligence and emotional instability, although the latter difference is small and insignificant." "It is suggested that better preparation and more rigorous selection would account for much of this superiority."

TABLE XVII.—SUMMARY OF GARRETT'S DATA

	N	Grades		Thorndike		Social Intelligence		Personal inventory	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Jewish	82	60.93	14.05	88.04	8.53	115.32	11.40	22.29	8.34
Native	65	37.56	13.95	79.04	11.23	106.94	12.19	19.00	7.44

GENERAL SUMMARY AND CONCLUSIONS

A critical resume of twenty-three comparative studies on the intelligence of Jews and non-Jews is presented. There is no study available as yet against which serious objections cannot be lodged. The authors of the studies reviewed violated fundamental criteria of valid research in the field of racial mental differences. As regards the factor of socio-economic status, for instance, although it is generally agreed that no study of racial mental differences is conclusive unless this factor is ruled out, either experimentally or by statistical techniques, yet in only three of the studies reviewed was this done. Again, while it is generally agreed that in studying racial mental differences the factor of language handicap should be ruled out, yet in only one study was this factor adequately controlled. It appears, therefore, that the results are highly tentative. The body of objective facts concerning the "psychology of the Jew" as it is reflected in differences in intelligence is quantitatively meager and qualitatively highly inferior. Investigators, however, are apparently satisfied with the results at hand, as no major comparative study of Jewish and non-Jewish intelligence has been reported since 1930.

Because of the fact that the studies reviewed are not of equal scientific value, a great many of them decidedly unsatisfactory, a

tabulation of the number of studies that found one group to be superior to the other would not only be masking the actual situation, but might also be misleading. The following conclusions are therefore the sum total of the impressions gained by the writer after a thorough analysis of the literature on the differences in intelligence between Jews and non-Jews:

(1) The Jewish children in Great Britain and the United States were found to be superior or at least equal in general intelligence to non-Jewish children of similar socio-economic status.

(2) The Jewish children were found, in most cases, to be superior to the children of other foreign-born racial or national groups, notably those of Italian and Central and Eastern European stock.

(3) Differences in intelligence were found between Jewish groups coming from different countries in Europe.

(4) The Jewish college student was found to be definitely superior to the non-Jewish student. This may be attributed to the more rigorous selection of the Jewish college student.

(5) The distributions of the general intelligence of Jewish children were found, as a rule, to be more homogeneous than those of non-Jews.

(6) As to frequency of high and low scores, the Jewish children compared favorably with non-Jewish native white in the United States.

(7) As to the difference between Jewish and non-Jewish children in non-verbal intelligence, there are not sufficient data to warrant conclusions.

(8) Nothing is definitely known of the differences between Jews and non-Jews in the rate of growth of intelligence.

(9) The problems of the differences between Jews and non-Jews in average general mental level, variability of the distributions frequency of mentally superior and inferior, growth of general intelli, gence and manual ability are still open to rigorous research.

REFERENCES

1. Maller, J. B.: "Studies in the Intelligence of Young Jews." *Jewish Edu.*, Vol. III, 1931, pp. 29–39.
   Rumyaneck, J.: "The comparative Psychology of Jews and non-Jews." *Brit. Jour. Psych.*, Vol. XXI, 1930–1931, pp. 404–426.
2. Pearson, K. and Moul, M.: "The Problem of Alien Immigration into Great Britain." *Annals of Eugen.*, Vol. I, 1925, p. 125.

3. Davies, M. and Hughes, A. G.: "An Investigation into the Comparative Intelligence and Attainments of Jewish and Non-Jewish School Children." *Brit. Jour. Psych.*, Vol. XVIII, 1927–1928, pp. 134–146.
4. Hughes, A. G.: "Jews and Gentiles." *Eugen. Rev.*, Vol. XX, 1928, pp. 89–94.
5. Winch, W. H.: "Christian and Jewish Children in East End Elementary Schools." *Brit. Jour. Psych.*, Vol. XX, 1930, pp. 261–273.
6. Weintrob, J. and Weintrob, R.: "The Influence of Environment on Mental Ability as shown by Binet-Simon Tests." *Jour. Edu. Psych.*, Vol. III, 1912, pp. 577–583.
7. Van Denberg, J. K.: *Causes of the Elimination of Students in Public Schools of New York City.* Teach. Coll. Contrib. to Edu., No. 47. New York: 1911, p. 38.
8. Judd, C. H.: *Measuring the Work of the Public Schools.* 1916, p. 290.
9. Brigham, C. C.: *A Study of American Intelligence.* Princeton University Press, 1923, p. 190.
10. Halpern, F. C.: *The Intelligence of Bilingual Jewish Girls.* M.A. Thesis, Columbia University, 1933.
11. Maller, J. B.: "Economic and Social Correlatives of School Progress in New York City." *Teach. Coll. Rec.*, Vol. XXXIV, 1933, pp. 655–670.
12. Hirsch, N. D. M.: "A Study of Natio-Racial Mental Differences." *Genet. Psych. Monog.*, Vol. I, 1926, pp. 233–406.
13. Brigham, C. C.: "Intelligence Tests of Immigrant Groups." *Psych. Rev.*, Vol. XXXVII, 1930, pp. 158–165.
14. Grier, N. M.: "Comparative Mentality of Jews and Gentiles." *Ped. Sem.*, Vol. XXV, 1918, pp. 432–433.
15. Murdock, K.: "Study of Race Differences in New York City." *Sch. and Soc.*, Vol. XI, 1920, pp. 147–150.
16. Jordan, R. H.: *Nationality and School Progress, a Study in Americanization.* Bloomington, Ill.: Public School Publishing Co., 1921, pp. 105.
17. Berry, C. S.: "Classification by Tests of Intelligence of 10,000 First-Grade Children." *Jour. Ed. Res.*, Vol. VI, 1922, pp. 185–203.
18. Bailey, P.: "A Contribution to the Mental Pathology of Races in the United States." *Ment. Hyg.*, Vol. VI, 1922, pp. 370–391.
19. Pintner, R. and Keller, R.: "Intelligence Tests of Foreign Children." *Jour. Ed. Psych.*, Vol. XIII, 1922, pp. 214–222.
20. Dearborn, F. W.: "Repeated Measurement of the Physical and Mental Development of School Children." *Sch. and Soc.*, Vol. XX, 1924, pp. 515–518.

    Dearborn, F. W.: *Intelligence Tests.* Houghton Mifflin Co., 1928, pp. 130–131.
21. Bere, M.: *A Comparative Study of the Mental Capacity of Children of Foreign Parentage.* Teach. Coll. Contrib. to Edu., No. 154. New York: 1924, pp. 105.
22. Feingold, G. A.: "Intelligence of the First Generation of Immigrant Groups." *Jour. Edu. Psych.*, Vol. XV, 1924, pp. 65–83.
23. Fukuda, T.: "A Survey of the Intelligence and Environment of School Children." *Am. Jour. Psych.*, Vol. XXXVI, 1925, pp. 124–139.
24. Terman, L. N.: *Genetic Studies of Genius. Vol. I. Mental and Physical Traits of a Thousand Gifted Children.* Stanford University Press, 1925, p. 56.

25. Luckey, B. M.: "Racial Differences in Mental Ability." *Sci. Month.*, Vol. XX, 1925, pp. 245–248.
26. Seago, D. W. and Koldin, T. S.: "A Comparative Study of the Mental Capacity of Sixth Grade Jewish and Italian Children." *Sch. and Soc.*, Vol. XXII, 1925, pp. 564–568.
27. Graham, V. T.: "The Intelligence of Italian and Jewish Children in the Habit Clinics of the Massachusetts Division of Mental Hygiene." *Jour. Abnor. and Soc. Psych.*, Vol. XX, 1925, pp. 271–277.
28. Rumyaneck: *Op. cit.*, p. 411.
29. Goodenough, F. L.: "Racial Differences in the Intelligence of School Children." *Jour. Exp. Psych.*, Vol. IX, 1926, pp. 388–397.
30. Cohen, I. L.: *The Intelligence of Jews as Compared with Non-Jews.* Columbus, Ohio: Ohio State University Press, 1927, pp. 43.
31. Hayes, E. C.: "Racial Groups in a University." *Sci. Month.*, Vol. XXVI, 1928, pp. 158–159.
32. Rigg, M.: "Some Further Data on the Language Handicap." *Jour. Edu. Psych.*, Vol. XIX, 1928, pp. 252–256.
33. Garrett, H. E.: "Jews and Others." *Person. Jour.*, Vol. VII, 1929, pp. 341–348.

INTRODUCTION TO INVERTED FACTOR ANALYSIS, WITH SOME APPLICATIONS TO STUDIES IN OREXIS[1]

W. STEPHENSON

INTRODUCTION

In a letter to *Nature*,[1] and in papers elsewhere,[2,3] I have described an inversion of previous factor theorems and the technique employing them. In the present paper I continue the exposition of this new approach to factor study, with special reference to the investigation of character, temperament, attitudes and orexis generally, at the same time taking the opportunity to introduce the technique to the wider circle of workers in America.

Present-day factor analysis is concerned with a selected population of persons (n) each of whom has been measured or scaled in m tests or traits etc. The $m(m-1)/2$ correlations for the m variables are subjected to factor analysis, using Spearman, Thurstone, or other theorems. I have shown* that this technique, and all its theorems, can be inverted. That is, N tests or other items can be measured or scaled for M persons, and the $M(M-1)/2$ correlations for these M variables can be factorized in terms of appropriate theorems. In current studies the variables are tests, etc., whereas in the inverted case the variables are persons. The population is a group of persons in the theorems of Spearman, Thurstone, and factorists in general; in the inverted technique the population is a group of tests or other units, although it can also be a group of persons.

At first sight this seems a very innocuous change, but it is one of some interest for experimental psychology. For convenience I have agreed with a suggestion made by Prof. G. H. Thomson[4] that r will stand for correlations between tests, etc. as variables, while Q will stand for correlation between *persons* as variables. The present-day factor technique may therefore be designated r-technique, and the inverted form, Q-technique.

SOLUTIONS OF A FUNDAMENTAL PROBLEM

The following matrix will further illustrate the nature of Q-technique, and will be referred to frequently in the sequel:

---

[1] From the Psychological Laboratory, University College, London.

* *Ibid.*[1,2,3]

		TESTS (or traits etc.)							
		1	2	3	4	...	...	...	N
PERSONS	a	${}_aX_1$	${}_aX_2$	${}_aX_3$	${}_aX_4$	...	...	...	${}_aX_N$
	b	${}_bX_1$	${}_bX_2$	${}_bX_3$	${}_bX_4$	...	...	...	${}_bX_N$
	c	...	...	etc.					
	.								
	.								
	.								
	M	${}_MX_1$	${}_MX_2$	${}_MX_3$	${}_MX_4$	...	...	...	${}_MX_N$

(1)

${}_aX_1$ represents the score given to an individual a for item 1, a similar connotation holding for the other X-scores. Spearman and the factorists who have drawn their inspiration from him have been concerned with the correlations between *columns* in the matrix. Q-technique, on the other hand, operates with the correlations between *rows* in the matrix. The data have to be standardized in rows, persons being variables and tests or traits the population.

Prof. G. H. Thomson* has also, independently, drawn attention to the correlation of rows in such a matrix. He reported a difficulty of a fundamental kind, that seemed to preclude correlation at the outset. The *units* in the rows, Professor Thomson indicated, are usually not homogeneous.

But in the case of estimates of character, temperament and the like, and for a very wide range of psychology there is no such difficulty. No matter how heterogeneous the units may be in rows, they can be rendered homogeneous *with respect to the individual, by arranging them in an order of some kind for their representativeness or significance for the individual concerned.* Thus, an individual may have twenty different personality traits, the units for which are markedly different; but it is possible to put these in order, those most characteristic of or significant for the individual being ranked or scored

* *Ibid.*

highly, while those of little relative significance are ranked or scored lowly. The value of this device for type psychology will be made clear in the sequel.

Or, quite generally, and particularly in the case of data involving units of *time* (as in intelligence test scores), scores can be standardized first in columns of the matrix (1), and then these standard values can again be standardized in rows.[2]

It is the object of this paper to amplify both these processes, and more particularly that for rendering heterogeneous traits, etc., homogeneous relative to the individual himself.

Before turning to this amplification it is helpful to give attention to the inverted form of the Spearman two-factor theorem, and to other features of *Q*-technique.

INVERTED FORM OF THE TWO-FACTOR THEOREM

In the two-factor theorem of *r*-technique, one begins by considering the scores for one *person* in several *tests*,[5] and correlations between the tests are arrived at. If there are four tests as variables (1, 2, 3 and 4), then the four can be considered as described by one general factor (g) and uncorrelated factors (s), if $r_{12} \cdot r_{34} = r_{13} \cdot r_{24} = r_{14} \cdot r_{23}$.

For *Q*-technique the theorem is as follows:

Let the scores for four *persons* (a, b, c, d) in test or trait 1 be represented as

$$
\begin{aligned}
{}_aX_1 &= m_a g_1 + n_a s_a \\
{}_bX_1 &= m_b g_1 + n_b s_b \\
{}_cX_1 &= m_c g_1 + n_c s_c \\
{}_dX_1 &= m_d g_1 + n_d s_d
\end{aligned}
$$

Variables are measured from their means, and the various s's are uncorrelated with each other and with g. g_1 is a constant for the *test* 1; the constants m and n are for the *persons* a to d respectively.

Then it may be proved as usual that if $Q_{ab} \cdot Q_{cd} = Q_{ac} \cdot Q_{bd} = Q_{ad} \cdot Q_{bc}$, the four variables may be considered as due to one general factor (g) and uncorrelated factors.

Also, $m_a^2 + n_a^2 = 1$, and similarly for any other person.

Further,

$$
\begin{aligned}
m_a &= Q_{ag} \\
n_a &= (1 - Q_{ag}^2)^{\frac{1}{2}}
\end{aligned}
\qquad (2)
$$

And,

$$Q_{ag}^2 = \frac{Q_{ab} \cdot Q_{ac} + \cdots}{Q_{bc} + \cdots} \qquad (3)$$

Thus, instead of calculating the saturations or loadings of *tests* with particular factors, by *Q*-technique we find such values for *persons*. Q_{ag} is the saturation or loading of person *a* with factor *g*. It is a constant for the individual, given uniquely by (3). It is not just an estimate from a regression line as measurements are in *r*-technique. Again, instead of *g* being a constant for a person it is now a constant for a particular test or trait etc. This may be estimated from a regression equation as usual.

IMPLICATIONS FOR TYPOLOGY

We have heard much, recently, about type psychology, especially from German workers. Stern, Jung, Jaensch, and Spranger, to mention only a few names at random, give ample evidence of a search for something other than current measurement in terms of universal factors and individual differences. It is usual to contrast *type* psychology and the psychology of individual differences. In the latter case, as I have said, the factors are considered to be universal, whereas in type psychology they need not be. But this is by no means the essential difference between the two psychological viewpoints, about which, however, I have yet to find any clear statement. The difference need not lie in the fact that for some type psychologists the whole subject of factors is absurd, on the ground that parts, such as factors, bear no relation or necessary resemblance to the *whole* personality. Vernon[6] has championed this gestalt notion, following the usual formula of Köhler and others, that the sum of factors cannot but give a distorted description of personality.

Persons who correlate together, or who are saturated with the same factor in *Q*-technique, obviously are alike in some respect. In a sense they approximate to a type. A factor can be regarded as a hypothetical "typical person," as shown by the system of correlations, and the saturation of a person with this factor is merely a measure of the extent to which he resembles this hypothetical type. By *Q*-technique we can determine types of persons; and Q_{ag} is an index of the extent to which the person approximates to the type. The saturation with a factor is at once an indication of type, and of the *amount* of it.

But the essence of type psychology, I believe, lies in the fact of similar *dispositions* of traits, factors, or other components which are descriptive of the individual. The rainbow has component colors, following each other in order. In the same way a person has components, of traits and the like, disposed in a certain fashion; persons with similar dispositions will be of a type, and any doubts about gestalt-qualitat do not arise. All this, however, will be given amplification elsewhere. Meanwhile an example may make the matter clear. I shall show in the sequel that a certified manic-depressive patient and a normal person may be alike in type for mood-condition: The order, frequency, and kinds of moods are similar and are similarly disposed for the two persons. That is all that type psychology is concerned about. The mood-conditions differ in *intensity* and no doubt also in duration, but this is no direct concern of type psychology. Only *r*-technique can determine how such persons differ in intensity or duration.

Thus, I fancy that *Q*-technique will clear away misunderstandings between type and individual psychologists. It will already be obvious to factorists that *Q*-technique holds in it the extension of factor methodology to type psychology. It has never before been possible to measure for types, and now a tool for that purpose is available.

GENERAL IMPLICATIONS

One or two general implications deserve mention in this introduction to *Q*-technique, which has much to say about its complementary *r*-technique.

Whereas in *r*-technique there is always considerable doubt about the nature of any factor, in *Q*-technique there is not the same uncertainty. In my experiment on color preferences,[7] a factor is merely a hypothetical order of preferences for colors, to which the persons of a type approximate. It should be clearly distinguished from a mere average.[3]

Again, in many respects *Q*-technique is logically and scientifically prior to *r*-technique. We shall have an example of this in the next section. I believe that *Q*-technique can remove certain assumptions now current in *r*-technique.

Factors in *Q*-technique will usually not be universal. One person may have one factor, and another have quite a different one. As Stern[13] has put it, not everyone is a pick-pocket, and the traits of a pick-

pocket may not be the same in system as those of the rest of us who are free from this manual deftness. But we can determine if need be *how many* persons are of one factor or type, and how many are of another. Thus, in an experiment on the preferences of children for the subjects of the school curriculum,[3] I found two main factors for both boys and girls. In factor *I* the scientific subjects were preferred most and the literary least; in factor *II* the preferences were the other way round, literary first and science least. Boys were in proportions four to one for factors *I* to *II* respectively; girls were in proportions two to one for factors *II* to *I* respectively. All of which, I fancy, is in accord with popular notions about the preferences of children for the subjects of the curriculum.

Finally, in *r*-technique we perform our experiments on large numbers of persons. Work on anything less than two or three hundred persons is scarcely thoroughly significant. In *Q*-technique the number of persons can be quite small. The significance of a correlation now depends upon the size of the population of test units in the matrix (1). To an extent that I have yet to examine, the significance also depends on the kind of selection of all such tests or test units. In any case we can now perform experiments on a few persons only, with a large number of test units. *Q*-technique is a tool for the laboratory, no less than for massive field work.[7]

APPLICATIONS TO OREXIS

After the above introduction I turn to describe certain applications of *Q*-technique to work on personality traits, character, temperament, and the like, in short, to the estimation of the orectic[8] characteristics of a person. I am to describe two applications, the first, in this section, will be familiar to some extent to all factorists; the second, dealt with in subsequent sections, I believe to be an important and new approach to personality study.

To focus attention we may consider a study such as that of Webb[9] on *Character and Intelligence*, the forerunner of all *r*-technique applications to the study of personality traits. In Webb's work all the persons of a population were estimated for traits such as perseverance, trustworthiness, and the like. Each trait was a variable, the population was the group of persons estimated, and the estimators were, so to speak, outside the picture, little account being taken of them, except to give the so-called reliability coefficients of the traits. These

variables were correlated, and analysis gave first *w*-factor, and then *c*-factor;[10] with further analysis I have no doubt that other factors will appear too.

Now by *Q*-technique we could investigate the estimators rather than the estimated persons. We could not hope thereby to learn much about these estimate persons, but only something about the estimators. It might have been sound policy, and I suggest that it will be in any future studies of this kind, to have this knowledge that *Q*-technique supplies at the outset, before attempting a study of the Webb kind.

I have just inferred that if two persons estimate a population for the same trait, the correlation between the two series of estimates is sometimes called the "reliability coefficient" of the trait. *Q*-technique, for such data, would begin with the factor analysis of the correlations of this kind for *many* persons. Each of, say, ten estimators would rate the population for each trait. For each of Webb's eighty traits there would thus be eighty correlation tables, each with ten variables, or as many variables as there are estimators. An example will make the matter clear.

Ten teachers were asked to estimate a group of one hundred children for the trait of *reliability*, using a 11-point scale. For my convenience all employed the same frequency distribution of scores, namely the following:

Score	10	9	8	7	6	5	4	3	2	1	0
Frequency	2	5	7	11	16	18	16	11	7	5	2

Thus we have a population of one hundred persons as the row in matrix (1), and ten persons as variables. The resulting correlations are shown in Table I.

TABLE I.—SHOWING *Q*-CORRELATIONS FOR PERSONS *A* TO *J*; ESTIMATING A ONE HUNDRED POPULATION OF CHILDREN FOR THE TRAIT OF RELIABILITY

Persons	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
A	..	74	67	59	42	34	30	47	15	29
B	..	..	60	65	37	25	45	51	31	34
C	..	..	..	53	45	20	18	26	20	35
D	..	..	..	..	20	41	32	37	29	43
E	..	..	..	..	..	26	28	40	17	36
F	..	..	..	..	..	..	48	57	63	42
G	..	..	..	..	..	..	..	71	47	40
H	..	..	..	..	..	..	..	..	43	57
I	..	..	..	..	..	..	..	..	..	49
J										

Persons *A* to *D*, it will be seen, correlated a mean amount 0.63 among themselves; persons *F* to *J* correlate 0.52 among themselves on the average; persons *A* to *D* correlate only 0.32 with persons *F* to *J*. Estimater *E* is about equally distributed between these two groups.

We can now factorize the table. Without going into detail, either two or three factors can be considered to describe the correlations. One (x) is specific to subjects *A* to *D*; the other (y) is specific to persons *F* to *J*; a third covers the whole table (z). Or, for a simpler pattern, one factor (z) may cover the whole table; and *either* set *A* to *D*, *or* set *F* to *J* (or *both*, as above), have in addition a factor (u), (or, if both, the factors are (x) and (y) respectively).

Examination shows that persons *A* to *D* have estimated the placid, submissive persons as highly reliable, whereas persons *F* to *J* have estimated the more active and direct signs of reliability. For *F* to *J* the population had to *show* reliability, and not be merely negative or neutral for it. That is, with respect to this trait there are two types of estimators.

In the first place, then, analysis has shown the need for a clearer definition of the trait to be rated. Secondly, we know that persons *F* to *J* are homogeneous for the trait *reliability*—they appear to be estimating the same kind of behaviour (and the same holds, of course, for persons *A* to *D*). This homogeneity, in any case, is a pre-requisite of work such as Webb's; really thorough work on traits should begin with a *Q*-technique analysis of the estimators. Thirdly, we know precisely which are the most *conforming* raters. The rater with the highest saturation with the particular factor most approximates to the hypothetical factor to which all are approaching. Thus, accepting the factor condition (u) for *F* to *J*, person *H* has the highest u-saturation, an amount 0.81, while person *J* has u-saturation 0.64, and is the least conforming. Fourthly, it is possible to weight each rater for the trait concerned in proportion to his saturation for the factor; or, one can so weight the different raters to give the greatest possible correlation with the factor.[1] Finally, with weighted values of this kind for each trait, we can proceed with a straightforward analysis of data by *r*-technique in the way followed, say, by Webb.

The above application of *Q*-technique is obviously a pre-requisite of any factor analysis of estimated traits by way of *r*-technique. Further, the saturation of a person with a factor gives a better account of the person and the trait he is estimating, than his crude reliability coefficient.[2] But the technique requires some control. The standard

to which all persons are tending to conform, the hypothetical factor such as u, is not necessarily the most valid standard. Should the estimators be ten farmers and one psychologist, the former will set the pattern for the correlations, and the latter would be uninfluential in that respect. The psychologist may not correlate at all with the farmers, and yet his ratings may well be the only valid ones. As in all matters of technique, conditions have to be controlled and subjected to a certain reason and circumspection. It would be a matter of some interest, incidentally, to find whether the same estimator tends always to supply the highest factor saturations for the traits rated.

ANALYSIS FOR TYPES

The above application does not arrive at knowledge about the individual differences of the population. The following applications, on the other hand, do so. The above process still involves traits as units of measurement, and therefore, according to gestalt notions, there must ensue an inevitable distortion of the *whole* individual. Man, we are reminded, is more than any sum of such traits. Again, it is suggested that some traits lack the universality demanded by r-technique: As has been mentioned, we are not all pick-pockets, and it is unscientific as well as libellous to suppose that we are. The technique now to be described does, I believe, counter these objections to a large extent. All measurement becomes a matter of the relative significance of traits *within the individual;* the individual is his own standard of reference; and one need only compare the separate person as a pattern or type, with other separate individuals.

There are two ways of approaching the examination of persons for their typology. The first, to be described immediately, concerns the pattern of traits for an individual; the second, referred to later, builds up the whole person again, so to speak, after he has been taken to pieces by current r-technique or, indeed, by Q-technique. The first is still analytical, yet with *whole* implications; the second is completely synthetical.

As an example of the former approach I select an item from a major study of Kretschmer's psychobiogram,[12] in which Q-technique has been applied to both normal and abnormal subjects.[14]

Kretschmer gives a schematized description of temperament, with a scale of *moods* as one aspect of this scheme.* I took from this scale the list of thirty moods given in Table II, and drew up a short but sufficient

* *Ibid.*, p. 224.

TABLE II.—SHOWING A "POPULATION" OF THIRTY MOODS (AFTER KRETSCHMER)

1. cheerful
2. cheerfully elated
3. irascible
4. elated
5. "sunny"
6. "comfortable"
7. quiet humour
8. quietly tender-hearted
9. cordial
10. good-humoured
11. affectionate
12. frequent alternations from pathetic to sad
13. sad
14. "sensitive plant"
15. humourless
16. serious
17. delicacy of feelings
18. irritable
19. excited
20. nervous
21. cool
22. severe
23. harsh
24. cold, apathetic
25. obtuse
26. moody indolence
27. ecstatic alternations
28. alternations from pathetic to annoyed
29. bad-tempered
30. heavily dejected

definition of each with the aid of the *Oxford Dictionary*, so that there could be little doubt about what the moods meant. The thirty moods, so defined, were printed on separate sorting cards. The object is to sort these cards to a given frequency distribution, in the order of representativeness or significance for the person under examination. Thus, as applied to a mental hospital, the psychiatrist sorts the cards so as to be characteristic of the patient's mood-condition over a period of, say, the past three months. The required frequency distribution may be as follows:

Score	10	9	8	7	6	5	4	3	2	1	0
Frequency	1	2	3	3	4	4	4	3	3	2	1

If mood six, say, is the most characteristic, for prevalency and intensity, of the patient, it is given a score ten; if mood sixteen should be least characteristic, it is given score zero. The next two most characteristic of the patient are given scores nine each; the two next least characteristic get score one each, and so on until all the moods have been given a score. Precisely the same procedure is followed if one performs the estimation on oneself.

The moods are now the population in matrix (1), whereas the patients are the variables. Several psychiatrists could estimate the same patients for these moods, as well as for other features of temperament, etc., and all that has been said in the previous section for a trait like *reliability* will apply equally here. The estimator may note those moods which are not applicable to the patient at all; later only those patients with the same list of moods need be compared.

As an example I give Table III, the correlations for sixteen persons as variables, with the above thirty moods as "population." Persons *a* to *e* inclusive are manic-depressive patients, and *h* to *l* inclusive are schizophrenia patients, at the Horton Mental Hospital, Epsom, England; persons *f*, *g*, *m*, *n*, *o*, and *p* are normal persons, in point of fact, my parents, wife, brother, sister, and myself.

TABLE III.—SHOWING *Q*-CORRELATIONS FOR SIXTEEN PERSONS *a* TO *o*, FOR A "POPULATION" OF THIRTY MOODS

Persons	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*o*	*p*
a	..	70	50	63	47	58	08	−43	−15	−50	−48	05	10	−33	−27	−36
b	..	..	34	45	54	48	36	−53	−35	−30	−41	−25	32	−10	−31	−15
c	..	..	..	66	71	27	13	13	14	14	−16	−13	−37	−65	−42	−38
d	..	..	..	..	67	35	24	−05	11	16	23	−17	−21	−71	−29	−54
e	..	..	..	..	..	41	32	07	−08	21	10	−20	−18	−43	−27	−22
f	..	..	..	..	..	..	43	−46	−33	−54	−39	−03	47	−05	22	34
g	..	..	..	..	..	..	..	−49	−66	−07	−51	−44	16	43	29	32
h	..	..	..	..	..	..	..	....	77	48	56	10	−54	−55	−49	−56
i	..	..	..	..	..	..	..	....	....	14	49	29	−32	−73	−67	−70
j	..	..	..	..	..	..	..	....	....	....	61	−14	−66	−25	−44	−41
k	..	..	..	..	..	..	..	....	....	....	....	17	−59	−47	−50	−35
l	..	..	..	..	..	..	..	....	....	....	....	....	17	−18	−41	−23
m	..	..	..	..	..	..	..	....	....	....	....	....	....	47	53	35
n	..	..	..	..	..	..	..	....	....	....	....	....	....	....	39	50
o	..	..	..	..	..	..	..	....	....	....	....	....	....	....	....	58
p																

Three groups of positive inter-correlations appear. The manic-depressives correlate positively among themselves a mean amount 0.57; the schizophrenes also correlate among themselves positively; so do the normal persons. Except for persons *f* and *g*, who belong partly to the manic-depressive group, the groups correlate either negatively, or scarcely at all, with each other.

The persons *f* and *g* supply the example I have already mentioned, in which certified manic-depressives and certain normal persons are shown to be alike in *type* for mood-condition, although of course vastly different in *intensities* and *durations* of such moods. I would repeat here that in the above example the moods have been compared for an individual, with respect to himself. There is no direct comparison between one person and another for either absolute intensity or duration of any of the moods. To compare one person with another for

these dimensions is a matter for *r*-technique, and for *r*-technique alone. *Q*-technique gives the type; and measures it; *r*-technique is concerned with quite a different matter.

These *Q*-correlations can be factorized. I find that I can get along very well with the simple two-factor theorem of Spearman; it satisfies common sense requirements. Thus, if we are interested in the manic-depressive syndrome as a type, we can use persons *a* to *e* as "reference values" for such a type, and determine any other person's factor saturation for this type, as well as the saturation of the "reference value" persons themselves. Similarly the schizophrenes might be used as "reference values"[5] for this broad schizophrene type—excluding person *l*, who happens to be a paranoia case whereas the others are straightforward dementia praecox cases. Or, the normal persons might be used as "reference values," but scarcely with the same insightfulness in the above case, because of the correlation of *f* and *g* with the manic-depressive type.

TABLE IV.—SHOWING SATURATIONS WITH (1) MANIC-DEPRESSIVE, AND (2) SCHIZOPHRENE, TYPES AS "REFERENCE VALUES"

Persons	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*
(1)	0.77	0.64	0.72	0.77	0.80	0.70	0.29	−0.12
(2)	−0.53	−0.56	0.02	0.12	0.02	−0.60	−0.58	0.93

Persons	*i*	*j*	*k*	*l*	*m*	*n*	*o*	*p*
(1)	−0.02	−0.08	−0.05	−0.17	−0.17	−0.57	−0.41	−0.42
(2)	0.60	0.50	0.81	0.02	−0.73	−0.69	−0.73	−0.70

In Table IV I give the common-factor saturations supplied by equation (3) (p. 356), for two of the above hypotheses. Row (1) shows the saturation of each person with the manic-depressive type, persons *a* to *e* being the "reference values." Row (2) shows the corresponding values for the schizophrene type, persons *h*, *i*, *j*, and *k* being the "reference values." Persons *f* and *g*, we see, have saturations 0.70 and 0.29 respectively with the manic-depressive type, and all the normals are negatively saturated with the schizophrene type.

But I am not concerned to give a detailed analysis of such a table. I have postulated very broad factors, but even within each divergences

can be noticed. Persons *a* and *b*, for example, are more manic than depressive, while *c*, *d*, and *e* are more depressive than manic. All such broad factors will sub-divide into lesser components with more and more analysis. I would refer only to one or two matters. First, the types depend markedly upon the kind of population (tests, etc.) selected—there will be no longer any excuse for the haphazard throwing together of tests etc., for the virtue of such populations will lie in the quality of the psychology that the experimentalist can put into them. Second, the correlational calculations become burdensome if large numbers of persons have to be worked upon. Forty-nine hundred fifty correlations could be calculated for a table involving one hundred persons as variables. But by using identical frequency distributions the burden is eased; and, usually, it will only be necessary to work with pre-arranged persons as "reference values." Thus, ten typical manic-depressives might be used as standard "reference values" for measuring all other persons for this type. Here again we open up a completely new field of measurements, a technique the very opposite of current "norms" and "normalizing" in terms of huge populations of persons. I hope someday to evolve an intelligence test, with a few persons as such standards, and in terms of whom all other persons might be measured!

Finally, the above *Q*-technique can be used almost immediately on introversion-extraversion rating schemes and the like. In fact it will, I think, replace much of present-day analysis of such personality schedules.

SYNTHESIS OF FACTORS

As another application of *Q*-technique there is the following. Let any number of persons be measured by *r*-technique for any list of traits or other variables, such as Spearman cognitive *g*, *p* (perseveration), *w* (will), *c* (temperament), etc. Let these be standardized in *columns* of the matrix (1) (page 354): That is, we have standard measures of these factors or traits.

Now use these standard measurements in turn as *population*, with the persons as variables, and let the rows of the matrix be correlated. I shall be surprised if factors do not emerge. Each again will represent a type of person—one could so depict the common-sense type of person, with much *w* and less *g*, and so forth.

Theoretically, in any case, *r*-technique took the person to pieces, but *Q*-technique can offer to put him together again.

CONCLUSION AND SUMMARY

In the above pages I give an introduction to an inversion of previous factor analysis and technique, and supply examples of its applications. In this new technique (*Q*-technique), tests, traits, etc., comprise the population, and persons are themselves the variables. In many respects this is the most logical use to which the correlational methods could have been put, and it is a matter of surprise that it should have been neglected for so long, in spite of the wealth of attention given to direct factor analysis, with its tests as variables, and persons as population.

A factor in *Q*-technique is a hypothetical person, to which all persons of a type approximate. The saturation or loading of a person with a factor is an index of this approximation—it is a measure of how closely the person approximates to the hypothetical "typical person."

Some general implications of this complete inversion of current factor analysis are briefly sketched, in particular the unequivocality of factors in *Q*-technique, the use of the technique to determine whether factors are universal or not, and its applications to experimentation on small numbers of persons only. One can measure, quite generally, all persons in terms of a small number of "reference value" persons as standards.

The more interesting applications to the study of personality are described, especially the connections with *type* psychology. In the past type psychology could not be subjected to factor analysis, and *Q*-technique is pre-eminently a method of isolating types and of measuring persons for their approximation to such types.

Two examples in this orectic sphere are supplied. One concerns the factor saturation of *estimators*, as distinct from the estimated population. This has wide applications, for example in the investigation of examiners of essays, drawings, etc. The other is the more important one, of value to type psychology generally. No matter how heterogeneous a number of traits may be, they can be rendered homogeneous with respect to the individual by arranging them in an order of some kind for their representativeness or importance for the individual concerned. An example is supplied, for a set of different *moods*, but others could as readily have been given for introversion, ascendency, and indeed for any personality or attitude schedule.

Finally, the use of *Q*-technique for the synthesis of factors is outlined. Thoughout, attention is drawn to features of type psychology as distinct from the psychology of individual differences.

REFERENCES

1. Stephenson, W.: "Technique of Factor Analysis." *Nature*, Vol. CXXXVI, 1935, p. 297.
2. Stephenson, W.: "The Inverted Factor Technique." *Brit. Journ. Psych.*, Vol. XXVIII, 1936, p. 344.
3. Stephenson, W.: "Application of *Q*-technique to Preferences for School Subjects." *Brit. Journ. of Educ. Psych.*, Vol. VI, 1936, p. 43.
4. Thomson, G. H.: "On Complete Families of Correlations, Etc." *Brit. Journ. Psych.*, Vol. XXVI, 1935, p. 63.
5. Spearman, C.: *The Abilities of Man* (*appendix*). London: Macmillan, 1927.
6. Vernon, P. E.: *Brit. Journ. Psych.*, Vol. XXIV, 1933, p. 167.
7. Stephenson, W.: "Correlating Persons Instead of Tests." *Character and Personality*, Vol. IV, 1935, p. 17.
8. Aveling, F.: *Personality and Will.* London: Nisbet, 1931.
9. Webb, E.: "Character and Intelligence." *Brit. Journ. Psych. Mono. Suppl.*, Vol. I, No. 3.
10. Garnett, M. C.: *Brit. Joun. Psych.*, Vol. X, 1919.
11. Wedeck, J.: *Factor Study of Psychological Ability.* Thesis, University London Library, 1933.
12. Kretschmer, E.: *A Text-book of Medical Psychology.* Oxford Univ. Press., 1934, trans. E. B. Strauss.
13. Stern, W.: "On the Nature and Structure of Character." *Character and Personality*, Vol. III, 1935, p. 270.
14. Stephenson, W. and Bartlett, R. J.: "Kretschmer's Psychobiogram." To be submitted to the *Journ. Medical Psychology.*

Note added June 1st 1936: A paper *The Foundations of Psychometry: Four Factor Systems* (submitted to Psychometrica) deals more completely with the matrix of persons and traits or other attributes of persons. It is there shown that the matrix can be standardized in four ways, supplying four factor systems, two involving *r* and two *Q* technique. The pair for *r* can be independent of the pair for *Q*.

In a paper *The Theory and Quantification of Psychological Types* (to be published shortly) the application of inverted factor analysis to type psychology is described in detail. The above is merely an introduction to a new and extensive field of psychological and factor work.

THE SCORING OF MATCHING-TESTS WITH UNEQUAL SERIES OF ITEMS

DWIGHT W. CHAPMAN

Harvard University

In connection with an experimental method in psychology, the writer has recently devoted a paper[1] to the solution of the following problem:

Let there be given a series of u X-items,

$$X_1, X_2, \ldots X_t, \ldots X_u,$$

and a series of t Y-items,

$$Y_1, Y_2, \ldots Y_t \qquad (t \leqq u),$$

such that correspondingly numbered items of the two series are true apposites of each other, and when paired constitute a correct matching. If now the Y-items are arranged at random against the X-items, what is the probability that such an arrangement will result in s correct matchings ($s = 0, 1, 2, \cdots t$)? Calling this probability $P_{(s)}$, we find

$$P_{(s)} = \frac{t!}{s!u!}\left[\frac{(u-s)!}{0!(t-s)!} - \frac{(u-s-1)!}{1!(t-s-1)!} + \frac{(u-s-2)!}{2!(t-s-2)!} - \cdots + (-1)^{t-s}\frac{(u-t)!}{(t-s)!0!}\right]. \quad \text{(I)}$$

where t is the number of items in the shorter series and u the number of items in the longer series.

If now we investigate the distribution of s (with relative frequencies given by the corresponding values of $P_{(s)}$), we find that the mean number (or chance expectancy) of correct matchings is

$$M_s = \frac{t}{u} \qquad \text{(II)}$$

and the standard deviation of the number of correct matchings is

$$\sigma_s = \sqrt{\frac{t}{u}\left[1 - \frac{t}{u} + \frac{t-1}{u-1}\right]}. \qquad \text{(III)}$$

[1] Chapman, D. W.: "The generalized problem of correct matchings." *Ann. Math. Statist.*, Vol. VI, 1935, pp. 85–95.

In the special case in which the two series of items are of equal length ($t = u$), we have $M_s = \sigma_s = 1$, as Zubin[1] has shown. But if the two series are of unequal length ($t < u$), both the expected number and standard deviation of correct matchings become less than unity, and we have in general

$$M_s < \sigma_s < 1.$$

I am indebted to Professor P. J. Rulon for the suggestion that equations (II) and (III) may be of value in the scoring of the Matching Test. Let an individual have produced s correct matchings in an attempt to pair t items with u other items. To regard his score simply as s is open to two objections: (*a*) That we should expect him to achieve a score of M_s by chance alone, and (*b*) that the difficulty of obtaining any given number of correct matchings in excess of M_s is a function of the relation between the magnitudes of t and u. It would seem reasonable, therefore, to obviate the first objection by expressing his score as a deviation from expectancy, *i.e.* as $s - M_s$, and the second difficulty by expressing this deviation as a multiple of the standard deviation of correct matchings (even though, since the distribution of s is enormously skewed, this deviation must not be regarded as having the same meaning as a standard deviate in a normal distribution).

We should have then, for the standard score on a matching test,

$$\text{Score} = \frac{s - M_s}{\sigma_s} = \frac{s - \frac{t}{u}}{\sqrt{\frac{t}{u}\left[1 - \frac{t}{u} + \frac{t-1}{u-1}\right]}}. \qquad \text{(IV)}$$

Thus if an individual produces four correct matchings by arranging six items against eight others, his score is

$$\frac{4 - 6/8}{\sqrt{6/8[1 - 6/8 + 5/7]}} = \frac{3.25}{.850} = 3.82,$$

whereas if he had produced the same number of correct matchings by arranging six items against ten others, his score would have been

$$\frac{4 - 6/10}{\sqrt{6/10[1 - 6/10 + 5/9]}} = \frac{3.40}{.757} = 4.49.$$

[1] Zubin, J.: "The chance element in matching tests." *This Journal*, Dec., 1933, pp. 674–681.

The following table of σ_e, computed from Formula (III) by Mr. H. P. Schoen, will prove convenient for the calculation of such scores.

TABLE OF THE STANDARD DEVIATION OF THE EXPECTED NUMBER OF CORRECT MATCHINGS WHEN t ITEMS ARE ARRANGED AGAINST u ITEMS

u \ t	1	2	3	4	5	6	7	8	9	10
2	.5000	1.0000								
3	.4714	.7454	1.0000							
4	.4330	.6455	.8292	1.0000						
5	.4000	.5831	.7349	.8718	1.0000					
6	.3727	.5375	.6708	.7888	.8975	1.0000				
7	.3499	.5017	.6227	.7284	.8248	.9147	1.0000			
8	.3307	.4725	.5844	.6814	.7691	.8504	.9270	1.0000		
9	.3143	.4479	.5528	.6431	.7244	.7993	.8696	.9362	1.0000	
10	.3000	.4269	.5260	.6110	.6872	.7572	.8226	.8844	.9434	1.0000
11	.2875	.4086	.5029	.5835	.6556	.7216	.7831	.8411	.8963	.9491
12	.2764	.3925	.4827	.5596	.6281	.6909	.7492	.8040	.8561	.9059
13	.2665	.3782	.4647	.5385	.6041	.6639	.7195	.7718	.8213	.8686
14	.2575	.3653	.4500	.5196	.5810	.6401	.6934	.7434	.7907	.8358
15	.2494	.3537	.4342	.5027	.5634	.6188	.6700	.7180	.7635	.8067

NOMOGRAPHS CONCERNING THE SPEARMAN-BROWN FORMULA AND RELATED FUNCTIONS[1]

J. NORMAN ARNOLD

Purdue University

AND

JACK W. DUNLAP

Fordham University

The numerous applications of the Spearman-Brown formula have resulted in several workers preparing tables giving the values of r_{nn} for the various values of r_{11} when the test has been lengthened "n" times. The generalized formula from which these tables have been computed is

$$r_{nn} = \frac{nr_{11}}{1 + (n - 1)r_{11}}$$

The inverse problem; namely, knowing the present reliability of a test how many times must the test be lengthened to get a specified reliability often arises; and an exact value for "n" is not given by the present tables. The first of the three following nomographs provides a rapid and accurate method for the exact solution of this problem.

In order to determine the number of times a test need be lengthened simply connect the values of r_{11} and r_{nn} on the appropriate scales by a straight edge, thread, or better with a straight line scratched on a strip of celluloid, and where the line crosses the "n" scale read the desired value. Obviously if r_{nn} is desired it is only necessary to connect the given values of "n" and r_{11} and read the value of r_{nn} where the line cuts the r_{nn} scale.

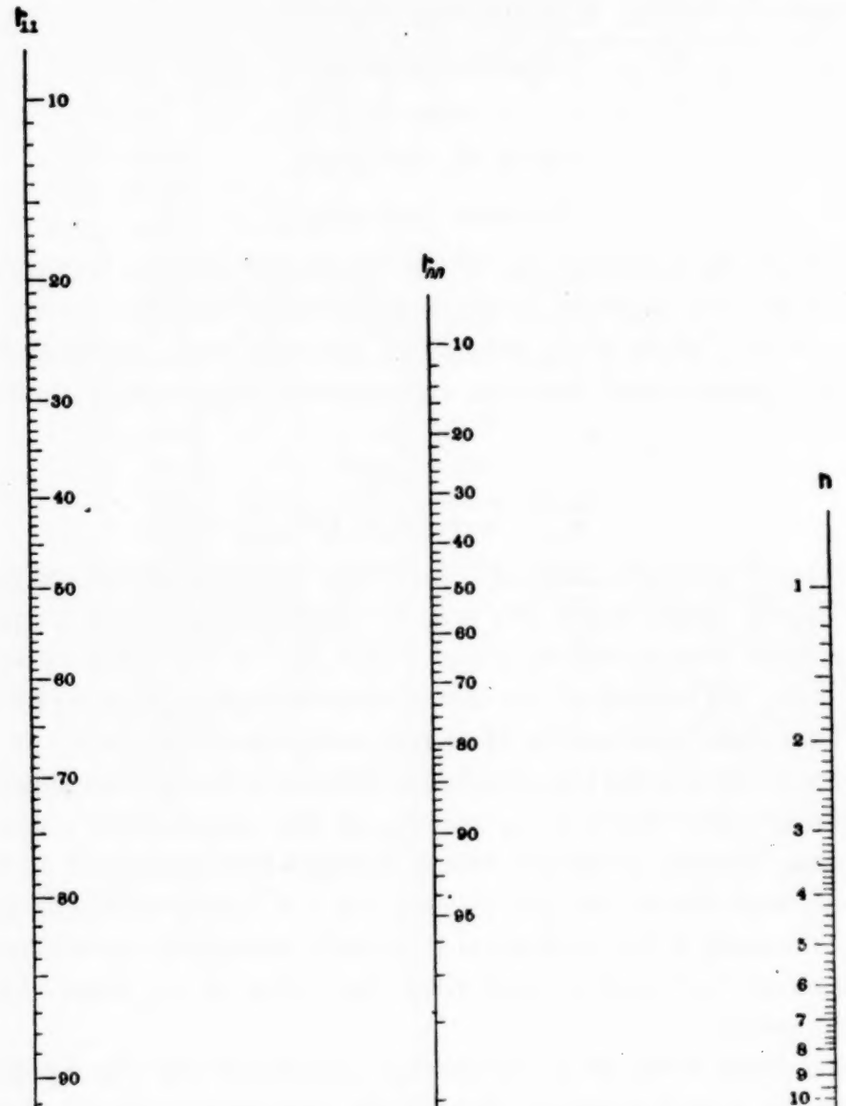
The sampling error of a correlation estimated by the Spearman-Brown formula was derived by Shen, 1924, and is given by the formula

$$\sigma_{r_{nn}} = \frac{n(1 - r_{11}^2)}{\sqrt{N}[1 + (n - 1)r_{11}]^2}$$

where r_{nn} = the estimated reliability
r_{11} = the observed reliability
N = the number of cases
n = the number of times the test was lengthened.

[1] The first of these nomographs was constructed by Dunlap, and the last two by Arnold.

The formula can be solved by means of the second nomograph. Given $r_{11} = 0.50$; $N = 100$; and $n = 4$. Locate the observed value r_{11}, here 0.50, on the r_{11} scale. Follow this line upwards to where it



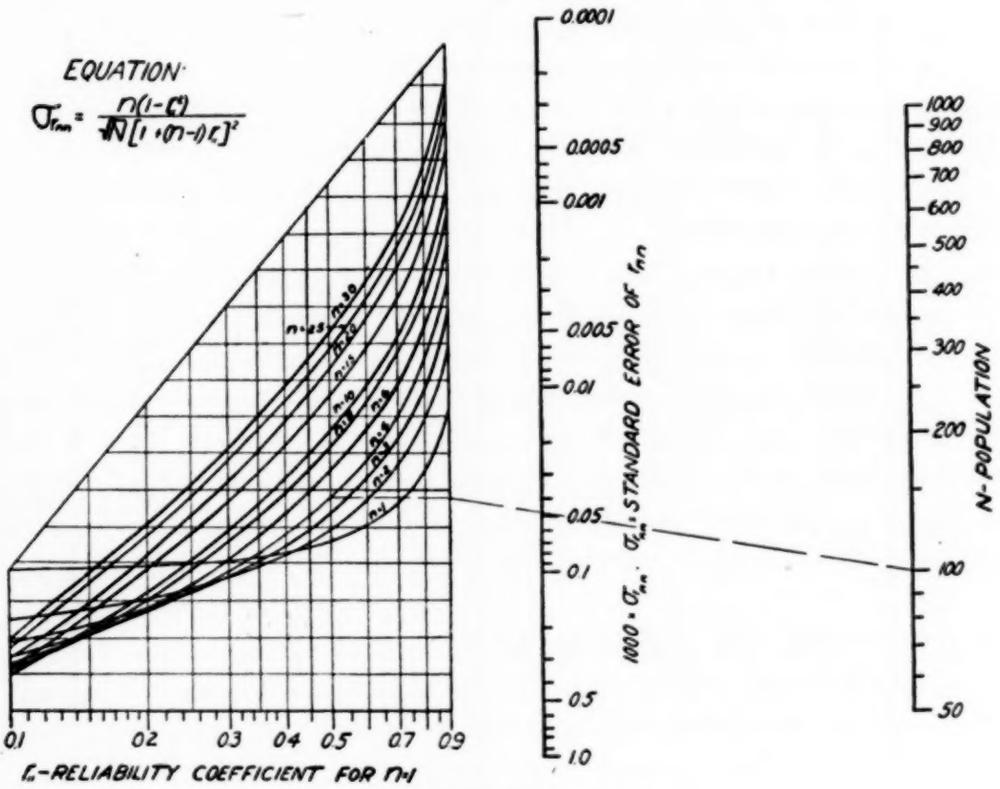
Nomograph I.

cuts the curve for 4, *i.e.*, "*n*." From this point project a line parallel to the base to intersect with the 0.9 line. Connect this point of intersection with one hundred, the size of the sample, on the *N* scale by

means of a straight line. Where the line cuts the $\sigma_{r_{nn}}$ scale read the desired value.

Often not only is "n" wanted, but it is desirable to have an idea of the magnitude of the sampling error of this estimated value, "n." Cureton, 1933, has developed the following formula for σ_n

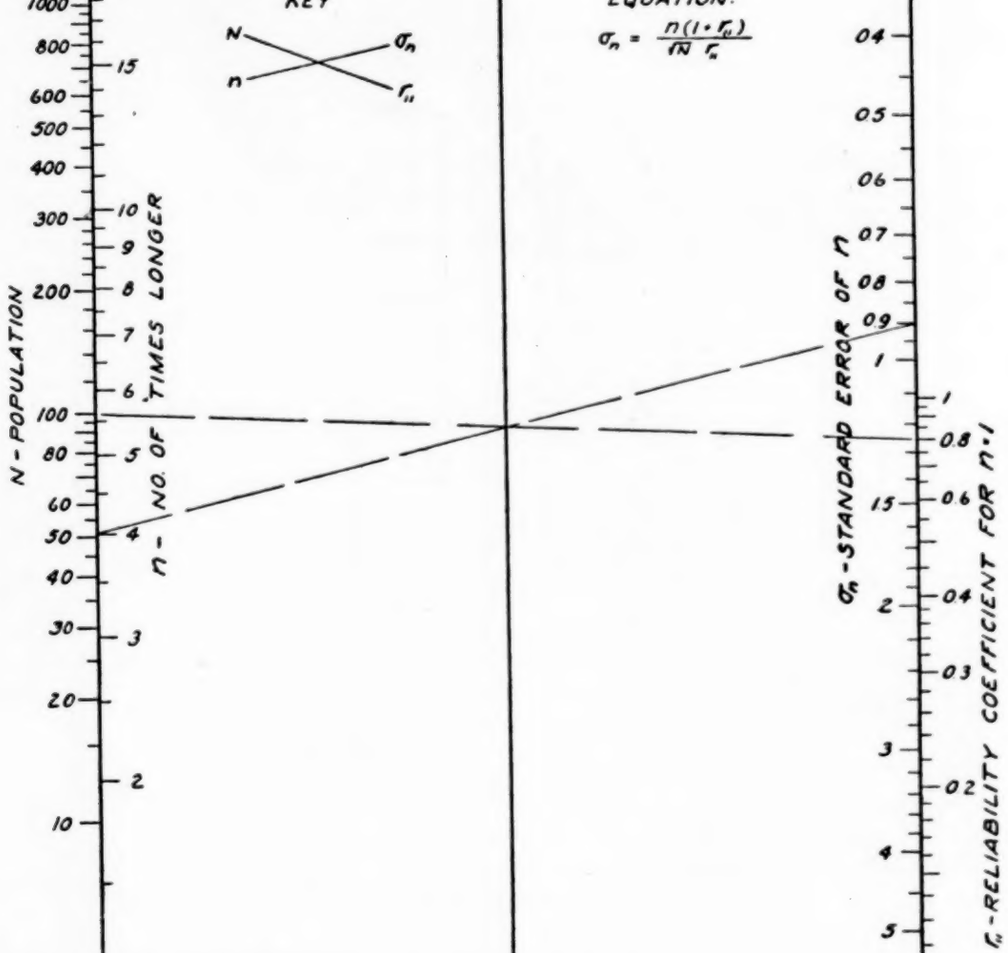
$$\sigma_n = \frac{n(1 + r_{11})}{r_{11}\sqrt{N}}$$



Nomograph II.

The solution of this formula is readily accomplished by means of the third nomograph. Given $N = 100$; $r_{11} = 0.80$; and "n" $= 4$. Connect one hundred on the N scale with 0.80 on the r_{11} scale by means of a hairline on a strip of celluloid. Pinning the celluloid strip to the middle or reference scale rotate the strip until the left-hand end of the hairline crosses the "n" scale at the estimated value, here 4. The desired value, σ_n, is read off the σ_n scale as 0.90.

The information obtained from these formulas and nomographs allows the test maker to determine the number of comparable items that must be added to get a desired reliability and also to estimate the excess number of items that need be included in order to allow for an underestimate due to sampling errors.



Nomograph III.

BIBLIOGRAPHY

Shen, Eugene: "The Standard Error of Certain Estimated Coefficients of Correlation." *J. Educ. Psychol.*, Vol. XV, 1924, pp. 462–465.

Cureton, E. E. and Dunlap, Jack W.: "A Nomograph for Estimating the Spearman-Brown Formula and for Computing its Probable Error." *J. Educ. Psychol.*, Vol. XXI, 1930, pp. 68–69. This is for the special case of "*n*" = 2.

Cureton, E. E.: "The Standard Error of the Spearman-Brown Formula when Used to Estimate the Length of a Test Necessary to Achieve a Given Reliability." *J. Educ. Psychol.*, Vol. XXIV, 1933, pp. 305–306.

A FACTOR ANALYSIS OF PERSONALITY TRAITS TO UNDERLIE CHARACTER EDUCATION

C. H. McCLOY

State University of Iowa

Much of the research in the last two decades in the field of what has been called "character" had tended to conclude that character education should be almost entirely specific. It has frequently found that training in one type of character responses or habits seem to have little effect upon character responses in other types of situations.

A careful analysis shows that the evidence supporting these claims is not entirely convincing. Some of this evidence is based upon experiments in the transfer of training in fields other than character education. Many such experiments have made assumptions which in the light of the facts seem difficult to justify, such as assuming that pupils who have been exposed to certain school subjects have, therefore, learned much concerning those subjects. Recent studies[1] have demonstrated that a very large percentage of pupils possess a high degree of natural immunity to learning, and escape almost entirely any infection with the virus thereof. Not having learned anything significant in one of the subjects included in a transfer experiment, it should not surprise us that such learning does not transfer to other subjects.

Other experiments more directly related to the field of character education have shown that boys who have been exposed to one type of so-called character education are no more superior in their resistance to certain specific forms of temptation than are other boys who have not been exposed to this training. On the basis of such experiments, the statement has sometimes been made that such character training is invalid and not effective insofar as any significant transfer to other fields is concerned. One plausible explanation might be that the individuals not subjected to this method of character education received just as effective training in their homes and in the schools, and it is entirely possible that the method of character training used upon this experimental group was so badly conceived in both form and content that little or nothing could be expected to result from it. It would seem to the writer that no real proof has been adduced to show that well conceived character training of a relatively general nature,

[1] Lindquist, E. E.: State University of Iowa, personal communication.

when organized upon lines which would be defensible from the standpoint of the laws of learning, will not be effective in the sense of transferring to other types of situations.

In attempting to simplify this problem it occurred to us to analyze a number of character traits to see whether any general or common elements might be found which could be used as the basis for a sound educational procedure in character education.

Two approaches to this seemed to be feasible, and both were tried. The first was an analysis by Spearman's two-factor method. Spearman's method is postulated upon the assumption that for any large number of traits there will be one common or general factor possessed in greater or less amount by all of these traits, and, in addition to this, each trait will be composed exclusively of a specific element. The general and specific elements together should account for all of the trait. This method is, of course, limited in application to situations in which it can validly be expected that only one common element will be found. From a careful study of the traits selected, it did not seem that this was apt to be true. The method was tried, however, on certain sub-groups of traits which seemed to have the possibility of possessing not more than one common element. These results were not particularly enlightening, and, since they have been reported elsewhere,[1] they will not be discussed here, except to state that in general their evidence supports that of the other method tried.

The second method used was that of Thurstone's multiple factor analysis. Thurstone's short or center of gravity method was selected and used in preference to other methods available for the same type of analysis.[2]

The Thurstone method of factor analysis enables one to analyze any given number of logically interrelated variables and to determine the amount of correlation each variable has with any number of elemental "factors" or components which are common to more than one of these variables. The process is carried out in successive steps

---

[1] McCloy, C. H.: "General Elements in Character." University of Iowa Studies in Physical Education, No. 2. Supplement to the *Research Quarterly of the American Physical Education Association*, Vol. VI, No. 3, October, 1935.

[2] Hotelling's method of analysis for common components and Kelley's multiple factor analysis involve much more work and the results give no more promise as to their accuracy and validity than those obtained from the Thurstone method. Hence, Thurstone's technique was selected as being as valid as the others and much more economical of time.

until the diminishing returns show that no more can be expected in the line of common factors. The results of this study are reported below.

Before proceeding further, it would perhaps be well to define our terms. The term "character" as used in this paper is not in any sense limited to the realm of morals. In this study it may be defined as the *determiners of conduct*, whether native or acquired. Probably most of these determiners are learned (*Cf.* Dewey, "the interpenetration of habits"), and since the educator can do nothing to change the heredity of an individual, he will do well to assume that all he can do about it is to influence this *learning basis of conduct.* This definition of character would include all of the social and individual qualities which the person exhibits in his day-by-day activities.

This study is based upon the inter-relationships of "traits" of character. It is a psychological commonplace to read that there is nothing in reality which corresponds to a generalized trait—that there are no such things as honesty, coöperation, or integrity that have any general application. The validity of this attack is acknowledged in part, but in part only; for large numbers of specific activities have in common general principles of conduct. Those relatively specific but interrelated behavior tendencies may be knotted together and carried by one verbal handle. The term "trait" will be used here quite unblushingly, and is defined as a term (denoting a concept) which brackets a number of specific or semi-specific habits or action-patterns having certain obvious things which fall, in general, under one heading, and which we have become accustomed to calling by one name. It is a sort of sub-gestalt, as it were, or integration of concepts, all of which are fashioned after one general pattern.

THE ORGANIZATION OF THE STUDY

Two sets of data were studied. These will be briefly described, and the results treated together.

(1) Thirty-one junior and senior students in the School of Physical Education of the George Williams College in Chicago were given rating sheets which had the names of forty-three traits listed thereon. The names of these traits are given in Table I. A short description or definition of each trait was appended and these were read to the group and explained. A method of rating in five categories was explained to the class and analyzed. Each student then prepared a list of other individuals whom he felt he knew well enough to rate with

TABLE I

The correlations given under columns I, II, III, and IV represent correlations with the common factors. These factors are identified by name at the head of the column. The column labeled "hypotenuse" indicates the relative distance away from the origin at which the trait is located when the factors are plotted in four dimensions, and enables the reader to determine how much of the trait is represented by the common elements and how much by specific elements. That accounted for by the general elements may be taken as the square of this hypotenuse, while the amount accounted for by specific elements and by errors of rating may be taken as $1 - \text{hypotenuse}^2$. It should be emphasized, however, that this error of rating is probably fairly large. Hence, it is probable that were the ratings more accurate, the amount of the common elements would be larger and the amount of the specific element relatively smaller. In view of the probable large error of rating, no valid attempt may be made to estimate the amount of specific elements.

	Social I	Positive action II	Individual III	Positive attitude IV	Hypotenuse
1. Adaptive	.33	.57	.13	.06	.67
2. Aggressiveness	.25	.78	.31	.10	.88
3. Alertness	.48	.18	.39	.17	.67
4. Athletic ability	.36	−.04	.74	.09	.82
5. Character	.77	.07	−.27	−.17	.83
6. Coöperation	.67	−.06	−.36	−.25	.80
7. Can discipline others	.46	.62	−.11	.03	.78
8. Conviction	.40	.43	.12	.09	.61
9. Cool-headed	.64	.08	.30	−.11	.72
10. Decisive	.42	.30	−.09	−.04	.53
11. Egotistical	−.06	.27	.19	.08	.35
12. Enthusiasm	.39	.40	−.32	−.06	.65
13. Energy	.21	.44	−.05	.78	.92
14. Fairness	.56	−.01	−.31	.09	.64
15. Fight	.51	.10	−.01	.04	.53
16. Good appearance	.45	.09	.33	.27	.63
17. Good follower	.60	−.10	−.41	.06	.74
18. Hard worker	.68	.00	.01	.31	.74
19. Health	.61	−.26	.14	.45	.81
20. Initiative	.36	.62	.12	.04	.73
21. Integrity	.75	−.03	−.35	.04	.83
22. Leadership	.71	.36	−.03	−.44	.91
23. Loyalty	.77	−.18	−.16	−.01	.80
24. Magnanimity	.62	.11	−.06	.06	.64
25. Modesty	.38	−.41	.00	−.36	.67
26. Moral courage	.73	−.10	.16	−.02	.76
27. Perseverance	.69	−.05	.00	.24	.73
28. Personality	.79	.00	.15	−.26	.84
29. Physical courage	.75	−.19	.00	.06	.78
30. Poise	.52	.30	.45	−.21	.78
31. Popularity	.60	.24	.36	−.22	.78
32. Resourcefulness	.75	.22	.28	−.19	.85
33. Respect for law	.64	.14	−.30	.21	.75
34. Respect for rights of others	.79	−.01	−.35	.02	.86
35. Self-confidence	.29	.63	.09	.24	.73
36. Self-control	.57	.14	.28	−.28	.71
37. Sociability	.75	.09	.31	−.08	.82
38. Self-sacrifice	.59	.26	−.12	−.02	.66
39. Self-denial	.61	.13	−.34	−.07	.71
40. Sportsmanship	.81	−.05	.09	.02	.82
41. Thoroughness	.77	.25	.00	.11	.81
42. Trustworthiness	.83	.00	−.17	.05	.84
43. Vitality	.73	.02	.18	.16	.77

accuracy, and each rated the three or four men whom he felt most competent to rate. Each person was also asked to rate himself. These students had lived together, in the same dormitory for the most part, for at least two years, and it is believed that the ratings were much more than usually valid.

The first step was to tabulate all ratings on each individual. It was found in general that the ratings agreed within one category.

Most frequently all of the ratings were in entire agreement, and, where they were not, usually all but one were in agreement. In cases where the ratings of others were evenly divided between two adjacent categories, the individual's own rating was used to break the tie. In general it was found that the individual's own rating agreed very closely with the majority opinion. In each case the modal rating was taken. All cases on whom ratings differed more than one category were eliminated from the study. No separate study on reliabilities was made.

A preliminary study upon a number of traits seemed to indicate an almost normal distribution of each of the traits within the group. It was assumed, therefore, that the different categories were almost evenly spaced, and all intercorrelations were performed by the product-moment method. It is possible that this assumption introduces some degree of inaccuracy. In the light of the crudity of the ratings, however, it is doubtful if any higher accuracy would have been introduced by a greater refinement of technique.

It will be noted that a number of traits which are not directly related to character have been introduced, as well as two general ones, "character" and "personality." These were introduced purely because of the curiosity of the writer to see where they would stand in such an analysis. They have been left in this study and incorporated in this report because of their suggestiveness for other studies in the field.

(2) In 1915, Webb reported a research entitled "Character and Intelligence,"[1] in which two groups of subjects were studied. The data from one of these groups were used here. The group was composed of one hundred ninety-four men with an average age of twenty-one years, all of them students in a teacher-training institution. Each was rated in the list of traits reported in Table II by two prefects who had had excellent opportunities to study the subjects for all of one semester, and who rated the students independently. No prefect rated more than twenty students. The details of the very excellent rating plan will be found in Webb's study. The ratings were made in seven categories, and were based on the theory that within such a homogeneous group, the ratings should follow a normal distribution. The intercorrelations were made by the product-moment method. Webb also corrected these correlations for attenuation, and obtained

[1] Webb, Edward: "Character and Intelligence." *British Journal of Psychology Monograph Supplement No.* 1, 1915.

TABLE II.—FACTOR LOADINGS FROM THE WEBB DATA

	Social I	Positive action II	Individual III	Positive attitude IV	Hypotenuse
1. Tendency to be cheerful	−.02	.19	−.50	.40	.72
3. Liability to extreme depression	−.12	−.16	.33	0.34	.55
4. Readiness to become angry	−.63	.33	.40	.31	.85
5. Readiness to recover from anger	.46	.07	−.54	.01	.72
9. Desire to excell in performance in which lies his chief interest	.29	.62	.27	.53	.85
10. Desire to impose will on others	−.53	.49	.30	.44	.83
11. Eagerness for admiration	−.61	.59	.20	.22	.88
12. Belief in own powers	−.28	.37	.41	.54	.80
13. Esteem of self as a whole	−.42	.47	.38	.43	.82
14. Offensive manifestation of self-esteem	−.61	.34	.40	.30	.84
15. Fondness for large social gatherings	−.47	.37	−.45	.33	.83
16. Fondness for small circle of friends	.47	.02	.23	.24	.57
17. Impulsive kindness	.30	.36	−.39	−.02	.61
18. Kindness on principle	.81	.26	−.31	.12	.91
19. Degree of corporate spirit	.19	.45	−.48	.26	.72
20. Trustworthiness	.86	.07	−.09	.11	.87
21. Conscientiousness	.87	−.01	.06	.04	.87
23. Readiness to accept sentiments of his associates	.30	−.07	−.32	−.42	.63
24. Desire to be liked by his associates	.13	.69	−.36	.13	.79
25. Wideness of influence	.44	.23	−.32	.48	.78
26. Intensity of his influence on his special intimates	.45	.23	−.10	.55	.77
27. Tact in getting on with people	.36	.00	−.50	.05	.63
28. Mental work on studies	.58	.22	.43	.31	.79
29. Mental work on pleasures	−.25	.31	−.43	.28	.65
31. Bodily activity on pleasures	−.29	.38	−.41	.27	.70
33. Tendency not to abandon tasks	.45	−.12	−.00	.42	.69
35. Quickness of apprehension	.19	.13	−.14	.62	.75
36. Profoundness of apprehension	.55	.00	.09	−.16	.83
38. Originality of ideas	.27	.08	.26	.66	.81
40. Physique	.14	.07	.01	.72	.83
44. Athletics	.24	.34	.34	.45	.62

a number of intercorrelations greater than one. We tried to use this table of "corrected" correlations, but at once encountered mathematical difficulties which threw considerable doubt on the validity of the assumptions made in correcting these correlations for attenuation; hence the study reported here used only the original raw zero order correlation coefficients. (See Table IV opposite page 26 in Webb's study.) In the present study, a number of the items included in Webb's study were omitted as irrelevant to our purpose.

MULTIPLE FACTOR ANALYSIS

In the use of any method of factor analysis the analysis is facilitated when not more than three common factors are found to be present.

This enables the factors isolated to be accurately plotted in three dimensions, usually from three orthogonal planes. Unfortunately, it was found that there was a fourth factor common to a number of the variables, thus throwing the problem into four dimensions. Since the original axes almost never correspond to the "real" axes, some method of rotation must be used. It can be easily demonstrated that planes or axes fitted purely by least squares seldom fit the facts. The fitting must be by a combination of both mathematics and logic. In the case of the study presented here, no especial difficulty was found in fitting the rotated planes, and the resulting "corrected" or "rotated" factor loadings (zero order correlations with the common factors or components) are given in Tables I and II.[1]

Further study of these factors enabled us to identify them beyond any reasonable doubt. The first common element was called "social qualities" as contrasted with "anti-social qualities." Social in this sense is to be distinguished from sociable and may be taken to mean those qualities which are beneficial to society or which are the qualities of a good citizen. There is a large group of such qualities, led by such traits as trustworthiness, sportsmanship, conscientiousness, respect for the rights of others, and thoroughness. There were few strongly anti-social traits, the nearest approaches being egotism, readiness to anger, offensive self-esteem, and eagerness for admiration. The second factor was called "positive action-tendency, or dominance," the opposite pole being "negative action-tendency, or submission." This is illustrated in the one extreme by aggressiveness and the desire to impose one's will on others, and in the other by modesty and suggestibility. The third column element is a little more difficult to name. It seems to be "individual qualities" as contrasted with the type of trait that causes individuals to merge with the group. The positive trait is the sort of thing which, in the terms of one of our popular magazines, causes one to "stand out from the crowd." It is the opposite of the sort of thing which causes the person to merge with other gregarious individuals and lose his

[1] The method of fitting the planes is given in Thurstone's *Vectors of Mind* (University of Chicago Press, 1935), Chapter IX. Three points were found on one of the four planes, two on a second, and one on a third, and the fourth is determined automatically if it can be assumed to be orthogonal to the other three, and there is some evidence, not presented here, that it is. In the Webb data the first plane had to be fitted by least squares from six points, since three satisfactory points could not be located.

individual indentity, which may, of course, be a desirable thing in many cases. This is illustrated by athletic ability, poise, self-confidence, self-esteem, and popularity on the positive side, and by the more social traits such as coöperation, readiness to recover from anger, cheerfulness, tact, and respect for the rights of others on the negative side. The fourth common element seems to be "positive attitudes," as contrasted with positive action-tendencies or dominance, and is illustrated by the traits of energy, health, originality, quickness of apprehension, intensity of influence on one's intimates, and self-confidence. At the other extreme will be found leadership, self-control, suggestibility, and modesty.[1]

The traits studied correlate in varying degrees with each of these four common elements. In none of them is absolutely everything accounted for by these four factors. It is highly probable, however, that the differences between the total correlation of these four elements and unity is accounted for more by inaccuracy and errors of rating than it is by certain specific elements. This belief is strengthened by the fact that the communalities are in many cases as large or larger than one would expect the reliabilities to be. Until such a study is made on much more valid ratings and measurements, the relative amounts of specifics and errors must remain in doubt. Since this study is concerned only with the common elements, and since the specific elements may be reasonably well deduced from the definition of the trait, these specifics need not concern us further here. Two of these traits studied in the American data were of a general nature, "character" and "personality." It will be seen that these traits correlate highly with the first or "social" factor.

Three others, "athletic ability," "health," and "vitality," were included to satisfy the curiosity of the writer. Athletic ability correlated more highly with the third or "individual qualities" factor than did any other trait. Owing to the fact that the American subjects were physical education majors, this fact might be discounted considerably, were it not for the fact that Webb found the same thing

---

[1] Harold M. Williams has recently published "A Factor Analysis of Berne's 'Social Behavior Patterns, in Young Children,'" (*Journal of Experimental Education*, Vol. IV, No. 2, December, 1935) in which thirty traits of pre-school children have been studied. He found two general factors which correspond to the first two found in the present studies on adults, but did not find further general factors. It is quite possible that it will be found that the third and fourth factors do not develop until later.

true with his students, though to a lesser extent. "Health" and "vitality," however, correlated surprisingly highly with the "social" factor, and the writer feels that this fact should at least be given thoughtful consideration in character education and mental hygiene programs.

A study of the vector directions of traits shows some interesting groupings or "constellations" of traits. One finds several such groups clustering around common centers.

1. Respect for the rights of others is surrounded by integrity, good follower, fairness, character, self-denial, trust-worthiness, and other passive but socially constructive virtues.

2. Aggressiveness is associated with initiative, conviction, and decisiveness.

3. Cool-headedness, resourcefulness, poise, self-control, sociability, and popularity are closely associated.

In the English data we have:

4. Eagerness for admiration, self-esteem, offensive self-esteem, desire to impose one's will on others, and readiness to anger in one cluster. This probably corresponds to the more generalized trait of egotism in the American data.

5. Cheerfulness, the tendencies to do both mental work and physical work on one's pleasures, and fondness for large social gatherings fall in one group.

6. Quickness and profoundness of apprehension, originality and persistence are closely grouped with wideness of influence and intensity of influence on one's intimates.

7. College professors may be interested to note that, except for some differences in the fourth factor, mental work on studies, profoundness of apprehension, and the tendency not to abandon tasks are closely grouped with fondness for a small circle of friends.

These groupings of traits suggest that perhaps we need to analyze each constellation of a better selected group of a larger number of traits in an attempt to determine the important sub-gestalten or generalized group patterns as an aid to clearer thinking in character education.

The next step in this analysis was to select the traits which correlated most highly with each of these four factors and their opposites, and to attempt to compute multiple correlations of these traits with the general factor. Since each of these factor loadings is comparable to a zero-order correlation with the common factor, it was further

assumed that the ordinary multiple correlation formula might be used, and multiple correlations were performed according to the usual techniques. A number of such lists of traits might be selected, each quite valid and valuable as a criterion of the general factor involved.

The next step was to reduce the number of traits rated as much as possible without reducing unduly the multiple correlations. These results are seen in Tables III to VI.

TABLE III.—MULTIPLE CORRELATIONS OF COMBINATIONS OF TRAITS CORRELATING HIGHEST WITH THE "SOCIAL" FACTOR

	Variables—Webb data		Variables—Y.M.C.A. data
1	Conscientiousness	1	Thoroughness
2	Trustworthiness	2	Sportsmanship
3	Kindness on principle	3	Respect for the rights of others
4	Mental work on studies	4	Loyalty
5	Profoundness of apprehension	5	Perseverance
		6	Resourcefulness
		7	Trustworthiness
		8	Physical courage

$R_{0.12345}$	.940	$R_{0.12345678}$	.974
$R_{0.1234}$	.939	$R_{0.1234567}$	.973
$R_{0.123}$	.928	$R_{0.123456}$	.973
$R_{0.12}$	.914	$R_{0.12345}$	.969
		$R_{0.1234}$	.957
		$R_{0.123}$	.936
		$R_{0.12}$	.874

TABLE IV.—MULTIPLE CORRELATIONS OF COMBINATIONS OF TRAITS CORRELATING HIGHEST WITH THE FACTOR OF "POSITIVE-ACTION TENDENCIES" OR "DOMINANCE"

	Variables—Webb data		Variables—Y.M.C.A. data
1	Desire to be liked by associates	1	Aggressiveness
2	Desire to excell in chief interest	2	Initiative
3	Desire to impose will on others	3	Decisive
4	Self-esteem	4	Self-confidence
5	Degree of corporate spirit	5	Can discipline others

$R_{0.12345}$	.865	$R_{0.12345}$	.843
$R_{0.1234}$	.862	$R_{0.123}$	.841
$R_{0.123}$	.835	$R_{0.12}$	.832
$R_{0,12}$	.806		

TABLE V.—MULTIPLE CORRELATIONS OF COMBINATIONS OF TRAITS CORRELATING HIGHEST WITH THE FACTOR OF "INDIVIDUAL QUALITIES"

	Variables—Webb data		Variables—Y.M.C.A. data
1	Mental work on studies	1	Poise
2	Belief in own powers	2	Popularity
3	Self-esteem	3	Alertness

$R_{0.123}$	.557	$R_{0.123}$	.620
$R_{0.12}$	.530	$R_{0.12}$	.609
		$R_{0.23}$	.564

MULTIPLE CORRELATIONS OF COMBINATIONS OF TRAITS WITH THE FACTOR OF "TENDS TO MERGE WITH THE GROUP"

	Variables—Webb data		Variables—Y.M.C.A. data
1	Readiness to recover from anger	1	Coöperation
2	Tendency to be cheerful	2	Integrity
3	Tact	3	Respect for rights of others
4	Degree of corporate spirit		
5	Fondness for large social gatherings		

$R_{0.12345}$	.775	$R_{0.123}$	.419
$R_{0.1234}$	.688		
$R_{0.123}$	.657		
$R_{0.12}$	.620		

TABLE VI.—MULTIPLE CORRELATIONS OF COMBINATIONS OF TRAITS CORRELATING HIGHEST WITH THE FACTOR OF "POSITIVE ATTITUDES"

	Variables—Webb data		Variables—Y.M.C.A. data
1	Physique	1	Energy
2	Originality	2	Health
3	Quickness of apprehension	3	Hard worker
4	Intensity of influence on special intimates	4	Modesty
5	Belief in own powers		
6	Desire to excell in chief interest		

$R_{0.123456}$	.830	$R_{0.1234}$	.870
$R_{0.12345}$	.821	$R_{0.123}$	.854
$R_{0.1234}$	.775	$R_{0.12}$	.838
$R_{0.123}$	.763	r_{01}	.785
$R_{0.12}$	.757		

CONCLUSIONS

From these results it would seem to be implied that a number of character qualities may be reduced to four important common or fundamental components which are overlaid by a number of specific elements. It would further seem that these four common components may be very accurately measured by a number of properly combined ratings, assuming, of course, that the individuals are well known to the raters and that enough raters are used to validate the results to within a reasonable degree of accuracy. We should therefore like to propose the basis of a tentative procedure for character education as related to the various fields of education.

1. In most situations these general factors are obscured or camouflaged by a mass of specific detail. It should be the endeavor of the educator to bring these common factors to light and, as it were, strip off their camouflage of specifics, and, by the use of what Thorndike has termed the principle of partial activity focused upon these prepotent elements, sensitize the individual to these general factors and cause him readily to recognize these fundamentals in all situations.

2. It should be recognized that there are different varieties of transfer processes. In many situations within one area of experience, common elements almost thrust themselves upon one's notice, as in the case of recognizing the similarity of the principles of teamwork in basketball and in a school project of clearing a field for football practice. Since such situations are "real" to the child, and are concrete and quite objective, they provide a good place for building fundamental concepts, emotionalizing them, and attaching verbal symbols to both the concepts and the feelings about them. Other situations which have less of the common elements apparent and are better concealed by specifics must be more carefully attacked by generalizing the principles so that their relationships are seen. Situations in which the only common elements are principles and verbal symbols (such as teamwork in baseball and teamwork in the government and in politics) should be attacked by leading the pupils first to generalize the principles and then to intellectualize those principles in such a way as to project them into areas of behavior as yet unmet in experience, in order that they may have general rules of behavior based upon intelligently analyzed concrete experiences which have involved these four fundamental factors of character.

Stated in another way, the process of generalization is simply carried one step farther than usual. In addition to bracketing together specific trait-actions having obvious common elements and labeling the result with a trait name, the groups of traits which are based upon more general principles are bracketed together and appropriately labeled with a usable and easily understood verbal symbol. The educator should endeavor to cause the pupil to progress from the specific to the trait, and from the trait to the common components of which groups of traits of each type are composed, and to attach significance and emotional drive to these principles of action.

3. It should then be the privilege of the instructor to take advantage of various situations as they arise in his instructional periods to point out constantly these common components and strengthen and guide this intellectualizing process until the application becomes semi-automatic.

4. Finally, it is suggested that the process of rating behavior in its educational aspects, particularly from the aspect of behavior qualities which the instructor desires to inculcate and strengthen, may be much simplified by refraining from rating a great mass of detail, and rating instead the individual's status in these fundamental qualities.

This paper makes no attempt to develop these proposals or theories into a system of character education. It is meant to be merely a report of a research, together with suggestions which would seem to spring from the results of the study.

NOTE ON THE COMPARATIVE EFFICACY OF LECTURE AND SOCIALIZED RECITATION METHOD *vs.* GROUP STUDY METHOD

WILLIAM H. BARNARD

Head Department of Education, Taylor University, Upland, Indiana

In the elementary school, the high school, and in college we have educational enthusiasts debating about "methods" of teaching. This may be a healthy sign, for it is not too much to say that practically all of our methods and educational theories fail to be convincing because they represent some person's philosophy of education. The philosophy to be compelling needs to be subjected to experimentation.

Below are the results of an experiment carried out by the writer at the college level with a group of forty-six sophomores studying educational psychology. These students were divided into equivalent groups primarily on the basis of the grades made in general psychology the preceding semesters,[1] and, secondarily, on the basis of scores made on the Thurstone Psychological Examination given at the beginning of their freshman year. The type of experiment was of the rotation variety which ran for a period of sixteen weeks.

After dividing the class into equivalent groups, the instructor designated the sections as Section I and Section II. Section I met with the instructor three times per week for eight weeks, whereas Section II met with the instructor once per week as a class and then met in groups consisting of from three to five students, twice per week. Each small group in Section II was then required to meet with the teacher for a very brief period—never more than ten minutes per week—such time being a part of their regular fifty-minute class period. Each of the Sections I and II, therefore, met for class or group work three fifty-minute periods per week. At the end of eight weeks the procedure with the sections was reversed, and Section II met regularly with the instructor, whereas Section I met once per week with the teacher and twice per week in group-study.

As a measure of achievement comprehensive new-type examinations of one hundred seventy-five responses were given at the end of each of the eight weeks' periods.

The data from the experiment are presented in the table on page 389 in the form of a frequency distribution. This is done so that we may see the total range of the scores and the frequency of each.

[1] Experiment carried on the second semester of the school year 1933–1934.

Analyzing the data we see that Section I was superior to Section II during both of the experimental periods. In case of the first eight-week period the mean difference was 3.9, whereas in the second eight-week period the mean difference was 2.1. When we apply the "Critical ratio" technique, that is, divide the difference in means in each case by the standard error of the means, we secure ratios of .50 and .30

Scores, with frequencies, made on test at end of first eight weeks				Scores, with frequencies, made on test at end of second eight weeks			
Section I		Section II		Section I		Section II	
Score	Frequency	Score	Frequency	Score	Frequency	Score	Frequency
154	1	147	3	144	2	132	3
147	0	140	0	138	0	126	1
140	1	133	0	132	1	120	1
133	0	126	0	126	0	114	2
126	2	119	3	120	2	108	0
119	2	112	3	114	0	102	1
112	3	105	0	108	2	96	3
105	0	98	3	102	4	92	2
98	5	91	1	96	2	84	3
91	3	84	4	90	2	78	2
84	3	77	1	84	2	72	1
77	0	70	0	78	2	66	2
70	0	63	2	72	0	60	1
63	1	56	2	66	1	54	0
56	1	49	0	60	1	48	1
49	1	42	1	54	1		
				48	1		
N = 23 Mean = 103.6 ± 5.0[1] SD = 23.8		N = 23 Mean = 99.7 ± 6.0[1] SD = 29.0		N = 23 Mean = 99.5 ± 5.3[1] SD = 25.3		N = 23 Mean = 97.4 ± 5.0[1] SD = 23.8	

[1] Standard deviation.

respectively for the first and second eight-week periods. These ratios are not statistically significant as they may be due to chance. We are only safe in saying that if the experiment was repeated, or a second experiment carried on under comparable conditions, the mean of Section I would stand sixty-nine chances in one hundred of being above the mean for Section II for the first eight-week period and sixty-two chances in one hundred in the second eight-week period. These

results are interesting when compared with those of a fairly closely related study by Hartman.[1]

The number of hours spent per week in study for each of the groups is not given, however, the difference was insignificant.

It seems to the writer that one fact is becoming increasingly clear—research and experimentation as to method in college as well as in the grades and in high school are vital and necessary in the discovery of truth and fact.

SUMMARY AND CONCLUSIONS

1. Section I was superior to Section II in both eight-week periods.
2. Despite the fact that Section I was superior to Section II in each of the experiment periods, the difference in achievement was not statistically significant as it may be due to chance.
3. Greater emphasis should be placed upon research and experimentation in determining method and the effectiveness of method.

---

[1] Hartman, G. W.: "Comparative Pupil Gains Under Individual Conference and Classroom Instruction." *Jrnl Ed. Psy.*, Vol. XXVI, 1935, pp. 367–372.

BOOK REVIEWS

GEORGE KINGSLEY ZIPF. *The Psycho-biology of Language.* Boston: Houghton Mifflin Company, 1935, pp. IX + 336.

Dr. Zipf has broken new ground in the study of language in several ways, especially by examining certain simple features of phonemes, morphemes, and words statistically. For example, he treats the well-known but unmeasured fact that often used words are short by a quantitative analysis of counts made by himself and others and concludes, first, that words become short by being used often, and, second, that there is "a tendency to maintain an equilibrium in the stream of speech between frequency on the one hand and what may tentatively be termed variety on the other," and that the nature of this equilibrium is such that over a wide range of frequencies the product of the number of words of a given frequency of occurrence by the square of that frequency is a constant.

Dr. Zipf has clearly in mind the general problems of why a language uses certain vocal acts and combinations thereof from the many it might use and why it uses them in the relative frequencies in which it does use them. He knows that there are natural causes for all these events and does not rely on mythical spirits of languages. He makes ingenious and daring use of the facts of relative frequency to explain many of these events.

It is not quite clear whether he regards the principle of equilibrium as a final explanation or as a stage in a progress toward such. I think that the latter is the sound attitude, and that the acts and satisfactions of speakers and listeners (in modern times, of writers and readers, also) are the forces producing the equilibria found by Dr. Zipf. I hope that he will apply his insight and resourcefulness to studying the influence of, for example, distinguishability of morphemes and of words, and "variety" as a value to the speaker-writer and as a value to the hearer-reader.

Language is the greatest human invention and the most important social tool. Its description and history have absorbed the attention of a long and large series of gifted scholars. Its dynamics and causation are equally worthy of study. The results which Dr. Zipf has attained by studying the frequencies of occurrence of various features of language should encourage the further use of quantitative methods. Relations, patterns and shifts of meaning offer promising fields.

E. L. THORNDIKE.

Teachers College, Columbia University.

J. B. STROUD. *Educational Psychology.* New York: The Macmillan Co., 1935, pp. XI + 490.

R. A. DAVIS. *Psychology of Learning; a Textbook in Educational Psychology.* New York: McGraw-Hill Book Co., 1935, pp. XII + 489.

The extensive researches in educational psychology made during the past generation have resulted recently in the production of several excellent basic texts. Among these must be numbered the two under review. Comparing these texts with the ones produced in the pre-scientific period, the change for the better is almost beyond belief.

Stroud's order of topics is somewhat unusual. Starting with the Instinct Hypothesis, he proceeds to the discussion of Emotion, the Physiological Basis of Mental Life, Elementary Statistical Concepts, Work and Fatigue, Learning, Factors Influencing the Speed of Learning, Retention and Recall, Intelligence: Its Nature and Origin, Mental Growth, Intelligence Testing, Educational Measurement, Development of Personality, Studies in Personality and Some Educational Problems. The last chapter is a rather poor acknowledgment of the need for applying psychology to the problems of Reading, Writing, Arithmetic and similar subjects. It is the weakest chapter in the book.

The work is written from the functional standpoint with a bias towards the environmental interpretation of behavior. It is well written, well printed and well bound and is a welcome addition to the stock of good educational psychologies.

Davis' book in the opinion of the reviewer is the better one. Its nineteen chapters deal with the Objectives and Methods of Educational Psychology, Bases of Improvement, Perception, Observation and Report, Neurological and Psychological Bases of Learning, Explanatory Principles of Learning, Improvement in Motor Learning, Improvement in Mental Learning, the Permanence of Learning, Transference and Interference in Learning, Influences Detrimental to Learning, Attitudes and Incentives, Attention and Interest, Guidance Techniques, Modes of Presentation, Techniques of Study, the Measurement of Achievement, Relationships of Physical and Mental Traits, Personality and Mental Hygiene. Such a comprehensive list of topics indicates that the book is badly named; the title and sub-title should be made to change places.

The book is very well documented; the footnotes indicate the extensive reading of the author, particularly in the literature found in the journals, yet in very few places does the author appear to have been overwhelmed by the bulk of his material. The chapters entitled Explanatory Principles of Learning and Transference and Interference in Learning are specially good.

Any teacher of educational psychology looking for a fairly elementary text, should select Stroud; those looking for a rather more scholarly and advanced text should certainly choose Davis.

PETER SANDIFORD.

University of Toronto.

JOSEPH H. GRIFFITHS. *The Psychology of Human Behavior.* New York: Farrar and Rinehart, 1935, pp. XXI + 515.

This is a text by a teacher who has been considerably influenced in choice of subject-matter, method of approach and emphasis by student reactions. Consequently the organization of the book departs somewhat from that of most elementary texts. The major sections are on what psychology is, learning, feeling and emotion, thinking and imagining, perceiving, and personality. Although the first chapter, which consists of an historical orientation to psychology, is a first attempt at this sort of thing in elementary texts, it is for the most part praiseworthy. This is followed by perhaps the best discussion of its kind on the contrast of the common sense view, literary interpretation and scientific method in the study of behavior.

In the analysis of learning, goal-seeking and insight are emphasized. A Gestalt hypothesis is favored in contrast to behaviorism which "has erected a scheme of interpretation of learning for which genuine scientific evidence, physiological or psychological, is lacking." Throughout the book the author evaluates contrasting views in a critical manner. Illustrations are cited from everyday life; practical applications are stressed. In general the author has succeeded very well in his aim to present psychology as a vital human interest and from a strictly scientific viewpoint. There are, however, a few instances of overemphasis and misinterpretation: (1) The description of trends toward insanity and excessive egocentricity in genius is perhaps overdone. (2) More up to date material than Luckiesh's is available on the legibility values of color combinations. (3) The description of "congenital word-blindness" as a factor in reading disability is wrong. (4) There is far from universal agreement that

"the chief factor in aiding localization is the phase difference of the sound in arriving at the two ears."

Nevertheless the author has produced a very readable, a thoroughly up-to-date and a well balanced text. It will appeal strongly to both student and instructor.

MILES A. TINKER.
University of Minnesota.

J. A. LONG, AND P. SANDIFORD. *The Validation of Test Items.* Toronto: University of Toronto Department of Educational Research Bulletin No. 3, 1935, pp. 126.

The authors define item validity in both its analytical and curricular aspects and outline a score of techniques which have been advocated for determining validity. These include proposals by Vincent, McCall, Barthelmess, Long, Bliss, Symonds, Kelley, Clark, Cook, and Henry. Evaluations of these proposals are followed by reviews of the experiments of Barthelmess, Cook, Lentz, and Smith, for each of which a summary of findings is given.

A preliminary experiment with ten different methods of validating physiography test items showed the bi-serial *r* method uniformly the best, and the so-called balance method uniformly the worst. As a result of this finding, and the earlier summary of experimental literature, the authors conclude that if only one validity method is to be used, that method should be bi-serial *r*, and that, of the shorter methods, the proposal of Clark yields the best results.

As a final study, five hundred six-choice vocabulary items were arranged in five forms of one hundred items each and administered to five hundred graduate students with instructions to guess at those answers which were in doubt. One hundred forty-four sets of papers were selected for study, an effort being made to secure a random sample of the whole. Elimination of eighteen faulty items left four hundred eighty-two, the score upon which was used as the criterion for validating each item separately. Sets of one hundred items were selected successively by each of thirteen techniques and scores on these sets were correlated with the criterion. Correlations were very high, with the better techniques differing very little from one another in the effectiveness with which they yielded a test composed of the best one hundred items. Methods which tend to assign validity values uninfluenced by item difficulty seemed to be inferior to methods which tend to assign higher validity values as items approach fifty per cent difficulty. Taking both effectiveness and ease of computation into

account, the Kelley technique seemed to be the preferable one among those which tend to select items without reference to item difficulty. Using a combination of several techniques yielded no better results than using the best technique by itself. P. J. RULON.
Harvard University.

EDWARD L. THORNDIKE. *Adult Interests.* New York: The Macmillan Co., 1935, pp. IX + 265.

Thorndike, in his *Adult Learning*, showed that age *per se*, did not reduce the efficiency of learning. In the present volume, he is concerned with the question of interests, which are probably of greater importance in adult education than in the education of children. A non-technical presentation of experiments aimed at the elucidation of the nature and behavior of adult interests form the basis for a discussion of adult education. These experiments show, in general, that, while under ordinary conditions the gross direction of interests does not change greatly during adulthood, it is possible to change them. The rules for changing old, or developing new, interests are the rules for learning in general. The author says, "The golden rules for the cultivation of an interest in A are: Connect that interest with A and reward the connection. Make wise use of contiguity, suggestion, imitation, or other causes to connect the interest with A. If the interest cannot be connected with A directly, connect it with $A + X$, and shift from $A + X$ to A, rewarding the connection with A whenever it occurs."

To this reviewer, the last three chapters of the book are of greater significance than the other nine. In these three chapters, Thorndike gives his views on adult education. That education of adults is possible and desirable has been amply demonstrated, but some educationists have felt that it should be offered to all and sundry. It is the old cry of democracy—equalization, but Thorndike reacts, "There is really neither justice, benevolence, nor reason in this doctrine of equalization." Perhaps the author's views are most adequately expressed in one paragraph which we quote.

> The resources available for adult education are, and always will be, limited. The amount to be spent should be so distributed that no change in the distribution would increase the common good. If by taking education away from some one person and giving it to some other the common good would be enhanced, the distribution is imperfect and should be changed. This implies that those who can use large amounts of education best should have first claim upon them.

As the experimental results are here presented in a very general fashion, it is impossible to evaluate them. But one would strongly recommend the reading of the last three chapters to all persons interested in adult education. C. M. LOUTTIT.

Indiana University.

ROBERT T. ROCK. *The Influence upon Learning of the Quantitative Variation of After-effects.* New York: Bureau of Publications, Teachers College, Columbia University, 1935, pp. IX + 78.

This is an experimental study to determine the relative influence of varying amounts of rewards and punishments on learning when motivation is uniformly high. Two of the experiments deal with code-learning and one with ball-tossing. The rewards and punishments consist of statements of "right" and "wrong" with their appropriate tokens which the subjects could convert into cash to keep when the experiments were concluded. The author points out that an important consideration in the three experiments was to note the direct rather than the indirect action of the after-effects of variations in rewards. He concludes that, while different punishments, as after-effects, varied slightly with the advantage in favor of the greater penalties, they were not nearly as effective in learning as the rewards. In general, he finds that punishments were about as likely to increase the recurrence of incorrect responses as rewards were to increase the recurrence of correct responses. While the satisfactions and annoyances created by the rewards and punishments employed in the study were similar to those of many life situations, the classroom teacher may not find them practical to use for motivating purposes. However, the results support an important hypothesis, and it is that annoyers of varying intensities have no such uniform effect on weakening the connections between situations and responses as satisfiers of different degrees of intensity have on strengthening the connections to which they are attached. ROBERT G. SIMPSON.

Carnegie Institute of Technology.

HARRY C. MCKOWN. *Character Education.* New York: McGraw-Hill, 1935, pp. XIV + 472.

This book gives a general survey of the field, emphasizing the importance of character and its relation to social conditions. It points out historical facts of moral instruction, various methods, and the opportunities, materials, and techniques available in class-

room, extra-curricular activities, and public and community agencies, with suggestions as to how the school, radio, theater, home, and church might be changed in the direction of providing conditions more favorable to character-training. There are numerous references to educational and general material. The scientific and theoretical treatment is somewhat superficial, consisting largely of quotations without thorough evaluation. Two examples will illustrate: (1) Objections to the trait approach in character education are quoted, followed by the unsubstantiated statement that the individual is surely not worse off because an attempt has been made to establish a general virtue. (2) Several pages are given to a discussion of heredity and environment, the whole point of which might as well have been stated in one or two sentences, since the materials, discussion, and quotations presented are inadequate to prove any point.

C. R. Garvey.

Carnegie Institute of Technology.

Guy Thomas Buswell. *How People Look at Pictures, A Study of the Psychology of Perception in Art.* Chicago: University of Chicago Press, 1935, pp. XV + 198.

This book should prove fascinating to three classes of readers—psychologists, artists, and laymen interested in art. Several important questions relative to the ocular behavior of viewing pictures are answered. The psychologist interested in perception will not be surprised at most of the findings, but the artist and art patron will be. Introspective evidence of the artist and of the interested layman as to the manner of looking at pictures is supported by the objective records in only a few instances. Revelations of this study are important to many artists and art teachers who explain certain principles of composition, *e. g.*, balance and rhythm, in terms of ocular behavior; because these objective records will not support such explanations.

A fortunate cluster of resources has made possible an investigation of unusual breadth and quality. The considerable expense was met principally by a sizeable grant from the Carnegie Foundation for the Advancement of Teaching. Acknowledgements are made of indebtedness to a personnel trained in the technique of recording eye-movements and to the Dean and Associate Dean of the Art School of the Art Institute of Chicago.

The author judiciously adapted the subject-matter to the interested public rather than to psychologists. Many pertinent questions the

latter would like to ask are unanswered in the interest of more palatable reading for the artist. One familiar with this field is surprised to find that it includes only a report of this one investigation, and at the author's failure to report earlier pertinent research. The book is simply and clearly written. Those trained in art will find it interesting from the first to the last page. Yet, as the author remarks, the most interesting and valuable parts of the book are the results shown in sixty-seven full page plates.

Results are presented and the text written to meet the objective stated in the closing sentence of the book. "This investigation is, therefore, presented to students of art in the hope that the evidence herein contained may have, when interpreted from their point of view, even greater significance than is apparent to the psychologist without special training in the field of art."

That the nature of the research is exploratory is evident in the plan of the investigation. One feels with the author that more data should be forthcoming on many propositions before other than tentative conclusions can be drawn.

ROY C. LANGFORD.
Kansas State College of Agriculture and Applied Science.

AUGUST AICHHORN. *Wayward Youth.* New York: Viking Press, 1935, pp. XIII + 236.

NORMAN FENTON. *The Delinquent Boy and the Correctional School.* Claremont, Calif. Claremont Colleges Guidance Center, 1935, pp. 182.

These two books afford an interesting contrast in methods of approach to the treatment problems of delinquency. Aichhorn, in a typical Teutonic manner, starts with a theory—which for him is true—and then shows how practice following the theory is successful in his institution. Fenton, in a typical American manner, presents detailed statistics about the boys in his institution and then rather tentatively suggests possible values in the rehabilitation program of the institution.

Aichhorn's book is a translation of the second edition of his *Verwahrloste Jugend.* His purpose is "to discuss the application of psychoanalysis to the treatment of delinquent youth." The discussion is an interpretive account of the author's own experience in the psychoanalytic treatment of delinquents. From this experience he says that there are three sorts of delinquency, *viz.* that due to an excess of love, to an excess of severity, or to a simultaneous combination

of the two. In any case the child is still showing infantile behavior inasmuch as he is behaving according to the pleasure principle. There can be no doubt, from the excellent case analyses presented, that a psychoanalytic type of therapy does relieve emotional conflicts and serves to correct the child's misbehavior. However, one wonders, while reading, whether the success is general with all cases or only with a few. Also there seems to be no account taken of the child where delinquencies are an integral part of the habit patterns built up from familial and community interactions. Does the child who has been taught and encouraged to steal by his parents have the same sort of emotional conflict as the basis of his delinquency as the children about whom the author is speaking? Can the feebleminded child be treated according to the plan suggested by Aichhorn? If the answers to these questions are negative, then it would be important to know what proportion of delinquents fall into these or similar categories.

Fenton describes, in a statistical vein, a group of four hundred consecutive admissions to the Whittier State School. In addition he presents a history of the school and a description of the methods of dealing with the boys from the time of admission until they are released from all supervision of the institution. The statistical analysis is very similar in its findings to most other studies of delinquents. There are the same low intelligence, broken homes, familial tensions, poor school achievement, and the rest. The guidance program of the school attempts to deal with each boy individually and to rehabilitate him to the end of his assuming a satisfactory place in the community when he leaves.

Contrast of these two monographs clearly indicates the lacks in each. Treatment as described by Fenton consists largely in re-education, building new habits, trying to give the child a new point of view. Little or nothing is proposed for dealing with the attitudes and conflicts of the child. Aichhorn emphasizes this later to an apparent exclusion of the former. Both are good; both are necessary; but probably not equally so for every case.

C. M. Louttit.
Indiana University.

Alice E. Watson. *Experimental Studies in the Psychology and Pedagogy of Spelling.* New York: Bureau of Publications, Teachers College, Columbia University, 1935, pp. III + 144.

The purpose of this study is to determine a more efficient technique of teaching spelling than is usually available for teachers to employ in

the classroom with the emphasis on the high-school level. It is a combination of experimental, empirical, and theoretical analyses dealing with spelling ability, vocabulary, errors, and methods. The author concludes that economy of learning to spell appears to be dependent more upon developing effective association techniques from valid procedures of linking, grouping, comparing, or contrasting words than upon the prevalent practice of attacking each word as an isolated unit. Larger thought-patterns built up through the integration of each specific detail of the pupil's spelling experience would therefore constitute the essence of such a technique. This idea appears to have promise of more value in learning to spell than is generally believed.

One of the most significant features of this study is the promising technique suggested for diagnostic and remedial work in spelling. The suggestions for remedial procedure are particularly valuable for the classroom teacher. In reference to determining and validating the general list of words from the composite vocabulary for high-school purposes, the author might have checked it against the Horn list of ten thousand words to determine their general use in written composition and correspondence. The Thorndike list being based on a reading vocabulary is not nearly as valid for this purpose as the Horn list. The study is a significant contribution to the theory and practice of teaching spelling. Its content in monograph form should be a valuable service handbook for teachers of spelling.

ROBERT G. SIMPSON.

Carnegie Institute of Technology.